

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 06:36:35 ; Search time 54 Seconds  
(without alignments)  
78.485 Million cell updates/sec

Title: US-09-700-643A-7

Perfect score: 82

Sequence: 1 CAWVASGRGIRPVGRX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001s.\*

5: Geneseq2002s.\*

6: Geneseq2003as.\*

7: Geneseq2003bs.\*

8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	98.8	15	3	AAY49296 19P2 liga
2	77	93.9	15	2	Aaw31400 Synthetic
3	77	93.9	15	2	Aaw97230 C-Termina
4	72	87.8	15	3	AAY49293 19P2 liga
5	72	87.8	20	2	Aaw31394 Human typ
6	72	87.8	20	2	Aaw97236 Human typ
7	72	87.8	20	3	AAY49294 19P2 liga
8	72	87.8	20	3	AAB10365 Human oxy
9	72	87.8	20	4	AAB90392 Prolactin
10	72	87.8	20	4	AG62534 Human CRH
11	72	87.8	20	5	Aae26404 Human prr
12	72	87.8	20	6	ABU60846 Peptide p
13	72	87.8	21	3	Aaw31395 Human typ
14	72	87.8	21	3	AAB10366 Human oxy
15	72	87.8	21	4	AG62535 Human CRH
16	72	87.8	21	6	ABU60847 Peptide p
17	72	87.8	22	2	Aaw31396 Human oxy
18	72	87.8	22	3	AG62536 Human CRH
19	72	87.8	22	4	AG62536 Peptide p
20	72	87.8	30	6	ABU60848 Human typ
21	72	87.8	31	2	AAY49299 19P2 liga
22	72	87.8	31	2	Aaw31391 Human typ
23	72	87.8	31	2	Aaw87615 Human 19P
24	72	87.8	31	2	Aaw97235 Human typ
25	72	87.8	31	3	AAY49291 19P2 liga

26	72	87.8	31	3	AAB10362 Human oxy
27	72	87.8	31	4	AAB90991 Prolactin
28	72	87.8	31	4	AAB90995 Prolactin
29	72	87.8	31	4	AG62531 Human CRH
30	72	87.8	31	5	AAE26401 Human prr
31	72	87.8	31	6	ABU60843 Peptide p
32	72	87.8	31	6	ABU60827 Peptide p
33	72	87.8	31	7	AGC71228 Human pep
34	72	87.8	32	2	Aaw31392 Human typ
35	72	87.8	32	3	AAB10363 Human oxy
36	72	87.8	32	4	AG62532 Human CRH
37	72	87.8	32	6	ABU60844 Peptide p
38	72	87.8	33	3	Aaw31393 Human typ
39	72	87.8	33	3	AAB10364 Human oxy
40	72	87.8	33	4	AG62533 Human CRH
41	72	87.8	33	6	ABU60845 Peptide p
42	72	87.8	87	2	Aaw31390 Human typ
43	72	87.8	87	2	Aaw97226 Human typ
44	72	87.8	87	3	AAB10361 Human oxy
45	72	87.8	87	4	AG62530 Human CRH

ALIGNMENTS

RESULT 1

AAY49296

ID AAY49296 standard; peptide; 15 AA.

XX

AC AAY49296;

XX 22-FEB-2000 (first entry)

XX 19P2 ligand peptide fragment.

DE Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;

KW pituitary; regulatory mechanism; central nervous system; pancreatic.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 15

XX Note= "C-terminal amide"

XX WO960112-A1.

XX 25-NOV-1999.

XX 20-MAY-1999; 99WO-JP002650.

XX 21-MAY-1998; 98JP-00140293.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying diseases related to ligand abnormality.

XX Disclosure; Page 27; 73pp; Japanese.

XX The invention provides a monoclonal antibody which has a specific reaction with the part peptide of the C-terminal of 19P2 ligand or its derivative. The antibodies can be used in diagnosis or to treat or prevent diseases associated with abnormality in the pituitary function regulatory mechanism (e.g. promotion of prolactin secretion), central nervous regulatory mechanism, and pancreatic function regulatory mechanism. The antibody-based immunoassay can also be applied in clarifying the physiological functions of the ligand and its derivative. Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand

SQ Sequence 15 AA;

Query Match 98.8%; Score 81; DB 3; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAWYASRGIRPVGR 14  
 |||||  
 DB 1 CAWYASRGIRPVGR 14

RESULT 2

AAW31400  
 ID AAW31400 standard; peptide; 15 AA.

XX AC AAW31400;

XX DT 06-APR-1998 (first entry)

XX DE Synthetic ligand 19P2-L31 peptide II.

XX KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;  
 KW pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent; antigen.

XX OS Synthetic.

XX PN WO9724436-A2.

XX PD 10-JUL-1997.

XX PF 26-DEC-1996; 96WO-JP003821.

XX PR 28-DEC-1995; 95JP-00343371.

XX PR 15-MAR-1996; 96JP-00059419.

XX PR 12-AUG-1996; 96JP-00211805.

XX PR 18-SEP-1996; 96JP-00246573.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;

XX PI Kitada C;

XX DR WPI; 1997-363672/33.

XX PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 XX function in the central nervous system, pancreas and pituitary gland.

XX PS Example 43; Page 151; 259pp; English.

XX CC This peptide contains the partial C-terminal sequence of the synthetic  
 CC ligand polypeptide 19P2-L31 which is capable of binding to a G protein-  
 CC coupled receptor protein. This peptide is used as an antigen to prepare  
 CC rabbit anti-bovine 19P2-L31 antibodies which are used in binding assays.  
 CC Pharmacological compositions containing this ligand may be used as a  
 CC pituitary function modulator, a central nervous system modulator or a  
 CC pancreatic function modulator. This ligand could have specific  
 CC applications as a prophylactic or therapeutic agent for dementia,  
 CC depression, hyperkinetic syndrome, disturbance of consciousness, anxiety  
 CC syndrome, schizophrenia, trauma, growth hormone secretory disease, hyper-  
 CC and polyphagia, hyperlipidaemia, hypercholesterolaemia,  
 CC hyperglycaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis,  
 CC renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis,  
 CC spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral  
 CC sclerosis, acute myocardial infarction, infertility, spinocerebellar  
 CC degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis  
 CC and/or oligogalactia. Assays can also be developed to screen compounds  
 CC which are capable of altering the binding activity of the ligand  
 CC affecting activation of the G protein-coupled receptor protein

SQ Sequence 15 AA;

Query Match 93.9%; Score 77; DB 2; Length 15;

Best Local Similarity 92.9%; Pred. No. 1.4e-06;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAWYASRGIRPVGR 14  
 |||||  
 DB 1 CAWYASRGIRPVGR 14

RESULT 3

AAW97230  
 ID AAW97230 standard; peptide; 15 AA.

XX AC AAW97230;

XX DT 06-MAY-1999 (first entry)

XX DE C-terminal ligand polypeptide derived antigen.

XX KW G protein-coupled receptor; GPCR; hypocoovarianism; gonocyst cocogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation; modulation;  
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; choriocarcinoma; hydatis mole;  
 KW irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia; prolactin secretion.

XX OS Synthetic.

XX PN WO9858962-A1.

XX PD 30-DEC-1998.

XX PF 22-JUN-1998; 98WO-JP002765.

XX PR 23-JUN-1997; 97JP-00165437.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Hinuma S, Kawamata Y, Fujii R, Matsumoto H;

XX DR WPI; 1999-105614/09.

XX PT Use of G protein-coupled receptor ligands - for modulating prolactin  
 XX secretion or placental function, e.g. for treating menopausal syndrome,  
 XX tumours, autoimmune disease or abnormal pregnancy.

XX PS Example 43; Page 123; 241pp; English.

XX CC AAW97229-31 represent a ligand polypeptide derived fragments used to  
 CC produce antibodies. The specification describes an agent for modulating  
 CC prolactin secretion which comprises a ligand polypeptide or a salt, for a  
 CC G protein-coupled receptor (GPCR) protein. The agents for promoting  
 CC prolactin secretion can be used for treating or preventing  
 CC hypocoovarianism, gonocyst cocogenesis, menopausal syndrome, euthyroid or  
 CC hypometabolism. They can be used for promoting lactation in a domestic  
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin  
 CC secretion can be used for treating or preventing pituitary adenomatosis,  
 CC brain tumour, emmenopathy, autoimmune disease, prolactinoma,  
 CC infertility, impotence, amenorrhea, galactorrhea, acromegaly, Chiari-  
 CC Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright syndrome,  
 CC lymphoma, Sheehan syndrome or dyszoospermia. The inhibitory agents can  
 CC also be used as contraceptives. The agents for modulating placental  
 CC function can be used for treating or preventing choriocarcinoma, hydatis  
 CC mole, irruption mole, abortion, unthrifty fetus, abnormal  
 CC saccharometabolism, abnormal lipidmetabolism or oxytocia

SQ Sequence 15 AA;

Query Match 93.9%; Score 77; DB 2; Length 15;  
 Best Local Similarity 92.9%; Pred. No. 1.4e-06;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 CAWYASRGIRPVGR 14
DB      1 CAWYASRGIRPVGR 14

RESULT 4
AAW49293
ID      AAW49293 standard; peptide; 15 AA.
XX
AC      AAW49293;
XX
DT      22-FEB-2000 (first entry)
XX
DE      19P2 ligand peptide fragment.
XX
KW      Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
KW      pituitary; regulatory mechanism; central nervous system; pancreatic.
XX
OS      Homo sapiens.
XX
FH      Key
FT      Modified-site
FT      Location/Qualifiers
FT      15
FT      /note= "C-terminal amide"
XX
FN      W09960112-A1.
XX
PD      25-NOV-1999.
XX
PF      20-MAY-1999; 99WO-JP002650.
XX
PR      21-MAY-1998; 98JP-00140293.
XX
FA      (TAKE ) TAKEDA CHEM IND LTD.
PI      Matsumoto H, Kitada C, Hinuma S;
XX
DR      WPI; 2000-039381/03.
XX
PT      New monoclonal antibodies, useful in diagnosis, as drugs and in studying
PT      diseases related to ligand abnormality.
XX
PS      Disclosure; Page 26; 73pp; Japanese.
XX
CC      The invention provides a monoclonal antibody which has a specific
CC      reaction with the part peptide of the C-terminal of 19P2 ligand or its
CC      derivative. The antibodies can be used in diagnosis or to treat or
CC      prevent diseases associated with abnormality in the pituitary function
CC      regulatory mechanism (e.g. promotion of prolactin secretion), central
CC      nervous regulatory mechanism, and pancreatic function regulatory
CC      mechanism. The antibody-based immunoassay can also be applied in
CC      clarifying the physiological functions of the ligand and its derivative.
CC      Sequences AAW49293-302 represent peptide fragments of the 19P2 ligand
XX
SQ      Sequence 15 AA;

Query Match      87.8%; Score 72; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      2 AWYASRGIRPVGR 14

RESULT 5
AAW31394
ID      AAW31394 standard; peptide; 20 AA.
XX
AC      AAW31394;
XX
DT      06-APR-1998 (first entry)
XX

```

```

DE      Human type G protein-coupled receptor ligand fragment 4.
XX
KW      G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
KW      pituitary; central nervous system; pancreas; prophylactic;
KW      therapeutic agent.
XX
OS      Homo sapiens.
XX
FN      W09724436-A2.
XX
PD      10-JUL-1997.
XX
PF      26-DEC-1996; 96WO-JP003821.
XX
PR      28-DEC-1995; 95JP-00343371.
PR      15-MAR-1996; 96JP-00059419.
PR      12-AUG-1996; 96JP-00211805.
PR      18-SEP-1996; 96JP-00246573.
XX
PA      (TAKE ) TAKEDA CHEM IND LTD.
XX
PI      Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;
PI      Kitada C;
XX
DR      WPI; 1997-363672/33.
DR      N-PSDB; AAV02431.
XX
PT      Ligand peptide for G protein-coupled receptor - acts by modulating
PT      function in the central nervous system, pancreas and pituitary gland.
XX
PS      Claim 2; Page 185; 258pp; English.
XX
CC      This sequence represents a peptide fragment from a novel human type
CC      ligand polypeptide corresponding to amino acid residues 34 to 53 of the
CC      sequence represented in AAW31390 and is used in an assay to monitor
CC      ligand binding to the G protein-coupled receptor protein. Pharmaceutical
CC      compositions containing this ligand may be used as a pituitary function
CC      modulator, a central nervous system modulator or a pancreatic function
CC      modulator. This ligand could have specific applications as a prophylactic
CC      or therapeutic agent for dementia, depression, hyperkinetic syndrome,
CC      disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
CC      growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
CC      hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
CC      cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
CC      rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
CC      amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
CC      spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
CC      osteoporosis and/or oligogalactia. Assays can also be developed to screen
CC      compounds which are capable of altering the binding activity of the
CC      ligand affecting activation of the G protein-coupled receptor protein
XX
SQ      Sequence 20 AA;

Query Match      87.8%; Score 72; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      7 AWYASRGIRPVGR 19

RESULT 6
AAW97236
ID      AAW97236 standard; peptide; 20 AA.
XX
AC      AAW97236;
XX
DT      06-MAY-1999 (first entry)
XX
DE      Human type ligand polypeptide fragment.
XX
KW      Rat type ligand; modulation; prolactin secretion;

```

KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cocogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dysospermia;  
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;  
 KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.

XX Homo sapiens.

OS  
 XX WO9858962-A1.

PN  
 XX 30-DEC-1998.

XX 22-JUN-1998; 98WO-JP002765.

XX 23-JUN-1997; 97JP-00165437.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Hinuma S, Kawamata Y, Fujii R, Matsumoto H;

XX WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal syndrome,  
 PT tumours, autoimmune disease or abnormal pregnancy.

PS Claim 3; Page 166; 241pp; English.

XX The present sequence represents a human type ligand fragment. It is used  
 CC in the course of the invention. The specification describes an agent for  
 CC modulating prolactin secretion which comprises a ligand polypeptide or a  
 CC salt, for a G protein-coupled receptor (GPCR) protein. The agents for  
 CC promoting prolactin secretion can be used for treating or preventing  
 CC hypovarianism, gonocyst cocogenesis, menopausal syndrome, euthyroid or  
 CC hypometabolism. They can be used for promoting lactation in a domestic  
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin  
 CC secretion can be used for treating or preventing pituitary adenomatosis,  
 CC brain tumour, emmenopathy, autoimmune disease, prolactinoma,  
 CC infertility, impotence, amenorrhea, galactorrhea, acromegaly, Chiari-  
 CC Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright syndrome,  
 CC lymphoma, Sheehan syndrome or dysospermia. The inhibitory agents can  
 CC also be used as contraceptives. The agents for modulating placental  
 CC function can be used for treating or preventing choriocarcinoma, hydatid  
 CC mole, interruption mole, abortion, unthrifty fetus, abnormal  
 CC saccharometabolism, abnormal lipidmetabolism or oxytocia

XX Sequence 20 AA;

Query Match 87.8%; Score 72; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWVASRGIRPVGR 14

DB 7 AWVASRGIRPVGR 19

RESULT 7

AA49294

ID AAY49294 standard; peptide; 20 AA.

XX AAY49294;

XX 22-FEB-2000 (first entry)

XX 19P2 ligand peptide fragment.

XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.

XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FT Modified-site 20  
 FT /note= "C-terminal amide"

XX WO9960112-A1.

XX 25-NOV-1999.

XX 20-MAY-1999; 99WO-JP002650.

XX 21-MAY-1998; 98JP-00140293.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying  
 PT diseases related to ligand abnormality.

PS Disclosure; Page 26; 73pp; Japanese.

XX The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function.  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in  
 CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand

XX Sequence 20 AA;

Query Match 87.8%; Score 72; DB 3; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWVASRGIRPVGR 14

DB 7 AWVASRGIRPVGR 19

RESULT 8

AAB10365

ID AAB10365 standard; peptide; 20 AA.

XX AAB10365;

XX 24-NOV-2000 (first entry)

XX Human oxytocin secretion promoting peptide SEQ ID NO: 35.

XX Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.

XX Homo sapiens.

XX WO200038704-A1.

XX 06-JUL-2000.

XX 22-DEC-1999; 99WO-JP007199.

XX 25-DEC-1998; 98JP-00369585.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;  
 XX WPI; 2000-452298/39.  
 XX Physiologically-active polypeptide recognized as ligand by G protein-  
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs  
 PT for diseases relating to oxytocin secretion and in veterinary medicine.  
 XX PS Disclosure; Page 63; 72pp; Japanese.  
 XX This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak  
 CC pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter  
 XX SQ Sequence 20 AA;  
 Query Match 87.8%; Score 72; DB 3; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRPVGR 14  
 Db 7 AWYASRGIRPVGR 19  
 RESULT 9  
 AAB90992  
 ID AAB90992 standard; peptide; 20 AA.  
 XX AAB90992;  
 XX 22-JUN-2001 (first entry)  
 XX Prolactin releasing peptide SEQ ID NO:166.  
 XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimide; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200069900-A2.  
 XX 23-NOV-2000.  
 XX 17-MAY-2000; 2000WO-US013576.  
 XX 17-MAY-1999; 99US-0134406P.  
 PR 10-SEP-1999; 99US-0153406P.  
 PR 15-OCT-1999; 99US-0159783P.  
 XX (CONJ-) CONJUCHEM INC.  
 XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
 XX WPI; 2001-112059/12.  
 XX Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity.  
 XX Disclosure; Page 244; 733pp; English.  
 XX The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a

CC reactive group (II) (e.g. succinimide and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity in  
 CC vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention  
 XX SQ Sequence 20 AA;  
 Query Match 87.8%; Score 72; DB 4; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRPVGR 14  
 Db 7 AWYASRGIRPVGR 19  
 RESULT 10  
 AAG62534  
 ID AAG62534 standard; peptide; 20 AA.  
 XX AAG62534;  
 XX 24-AUG-2001 (first entry)  
 XX Human CRH releasing protein related peptide SEQ ID NO: 35.  
 XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
 KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
 KW Addison's disease; adrenal gland hyperfunction; obesity.  
 XX Homo sapiens.  
 OS WO200135984-A1.  
 XX 25-MAY-2001.  
 XX 17-NOV-2000; 2000WO-JP008119.  
 XX 18-NOV-1999; 99JP-00327900.  
 PR 26-SEP-2000; 2000JP-00297073.  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX Kitada C, Matsumoto H, Hinuma S;  
 XX WPI; 2001-355552/37.  
 XX Use of G protein receptor ligand or peptide for controlling corticotropin  
 PT releasing hormone secretion.  
 XX Claim 4; Page 75; 90pp; Japanese.  
 XX The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
 CC present sequence is a peptide used in the exemplification of the  
 CC invention

```

XX SQ Sequence 20 AA;
Query Match 87.8%; Score 72; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14
DB 7 AWYASRGIRPVGR 19

RESULT 11
AAE26404
ID AAE26404 standard; peptide; 20 AA.
XX AC AAE26404;
XX DT 13-DEC-2002 (first entry)
XX DE Human PrRP-31 C-terminal peptide, PrRP-20.
XX KW Human; wakefulness; sleep disorder; prolactin releasing peptide receptor;
XX KW PrRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;
XX KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;
XX KW anticonvulsant.
XX OS Homo sapiens.
XX PN US2002037533-A1.
XX PD 28-MAR-2002.
XX PF 17-AUG-2001; 2001US-00932161.
XX PR 28-APR-2000; 2000US-00560915.
XX PA (CIVE/) CIVELLI O.
XX PA (LINS/) LIN S.
XX PI Civelli O, Lin S;
XX WPI; 2002-403931/43.
XX Screening for compounds useful for promoting wakefulness or sleep, and
XX PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep
XX PT apnea, comprises administering a prolactin releasing peptide agonist or
XX PT antagonist.
XX PS Disclosure; Page 25; 35pp; English.
XX CC The present invention relates to a method of screening for compounds for
XX CC promoting wakefulness or sleep in a mammal. The method involves
XX CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)
XX CC agonist or antagonist respectively and determining the ability of the
XX CC compound to promote wakefulness or sleep. The compounds identified from
XX CC the method are used in the therapy of epilepsy and other diseases
XX CC associated with absence seizures and in promoting wakefulness and sleep
XX CC in individuals having sleep disorders such as insomnia and narcolepsy.
XX CC PrRP receptor agonists may be used to treat common disorders which lead
XX CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and
XX CC psychogenic hypersomnia. PrRP receptor antagonists are useful for
XX CC promoting sleep and for treating insomnia such as adjustment sleep
XX CC disorder and psychophysiological insomnia. The present sequence is human
XX CC PrRP-31 C-terminal peptide, PrRP-20
XX SQ Sequence 20 AA;
Query Match 87.8%; Score 72; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14
DB 7 AWYASRGIRPVGR 19

RESULT 12
ABU60846
ID ABU60846 standard; peptide; 20 AA.
XX AC ABU60846;
XX DT 06-MAY-2003 (first entry)
XX DE Peptide production by gene recombination associated peptide #30.
XX KW Peptide production; low-molecular peptide; KISS-1; GPR8 ligand;
XX KW gene recombination.
XX OS Homo sapiens.
XX PN WO200292829-A1.
XX PD 21-NOV-2002.
XX PF 16-MAY-2002; 2002WO-JP004735.
XX PR 17-MAY-2001; 2001JP-00147341.
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX PI Nishimura O, Suenaga M, Ito T, Kitada C;
XX WPI; 2003-129302/12.
XX Process for producing peptides e.g. Kiss-1 peptide and GPR8 ligand for
XX PT subsequent applications by gene recombination technique through tandem
XX PT repeats to provide precursor protein with specific cleavage sites.
XX PS Disclosure; Page 68; 87pp; Japanese.
XX CC The invention describes a method of producing a peptide comprising the
XX CC excision of the N and C-terminals of a target peptide with enzymes or
XX CC chemically through the attached cleavage sites repeated by ligation in a
XX CC precursor protein. The method is for producing (low-molecular) peptides
XX CC e.g. Kiss-1 peptide and GPR8 ligand for subsequent applications by the
XX CC gene recombination technique through tandem repeats to provide a
XX CC precursor protein with specific cleavage sites. With this method, peptide
XX CC production can be carried out easily to provide large quantities of the
XX CC required peptides. This is the amino acid sequence of a peptide
XX CC associated with the peptide production method of the invention
XX SQ Sequence 20 AA;
Query Match 87.8%; Score 72; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14
DB 7 AWYASRGIRPVGR 19

RESULT 13
AAW31395
ID AAW31395 standard; peptide; 21 AA.
XX AC AAW31395;
XX DT 06-APR-1998 (first entry)
XX DE Human type G protein-coupled receptor ligand fragment 5.
XX KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
XX KW pituitary; central nervous system; pancreas; prophylactic;

```

KW therapeutic agent.  
 XX Homo sapiens.  
 OS WO9724436-A2,  
 XX PD 10-JUL-1997.  
 XX PF 26-DEC-1996; 96WO-JP003821.  
 XX PR 28-DEC-1995; 95JP-00343371.  
 XX PR 15-MAR-1996; 96JP-00059419.  
 XX PR 12-AUG-1996; 96JP-00211805.  
 XX PR 18-SEP-1996; 96JP-00246573.  
 XX PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;  
 PI Kitada C;  
 XX WPI; 1997-363672/33.  
 XX N-PSDB; AAV02432.  
 XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland.  
 XX Claim 2; Page 186; 258pp; English.  
 XX This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 54 of the  
 CC sequence represented in AA031390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a prophylactic  
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,  
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,  
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,  
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the  
 CC ligand affecting activation of the G protein-coupled receptor protein  
 XX Sequence 21 AA;  
 SQ  
 Query Match 87.8%; Score 72; DB 2; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRPVGR 14  
 DB |||||  
 7 AWYASRGIRPVGR 19  
 RESULT 14  
 AAB10366  
 ID AAB10366 standard; peptide; 21 AA.  
 XX AC AAB10366;  
 XX 24-NOV-2000 (first entry)  
 XX Human oxytocin secretion promoting peptide SEQ ID NO: 36.  
 DE Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 XX treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 XX veterinary medicine; milk production.  
 XX

OS Homo sapiens.  
 XX WO200038704-A1.  
 XX PD 06-JUL-2000.  
 XX PF 22-DEC-1999; 99WO-JP007199.  
 XX PR 25-DEC-1998; 98JP-00369585.  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX Matsumoto H, Kitada C, Hinuma S;  
 PI WPI; 2000-452298/39.  
 XX Physiologically-active polypeptide recognized as ligand by G protein-  
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs  
 PT for diseases relating to oxytocin secretion and in veterinary medicine.  
 XX Disclosure; Page 63; 72pp; Japanese.  
 XX This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak  
 CC pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter  
 XX Sequence 21 AA;  
 SQ  
 Query Match 87.8%; Score 72; DB 3; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRPVGR 14  
 DB |||||  
 7 AWYASRGIRPVGR 19  
 RESULT 15  
 AAG62535  
 ID AAG62535 standard; peptide; 21 AA.  
 XX AC AAG62535;  
 XX 24-AUG-2001 (first entry)  
 XX Human CRH releasing protein related peptide SEQ ID NO: 36.  
 DE Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
 XX analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
 KW Addison's disease; adrenal gland hyperfunction; obesity.  
 XX Homo sapiens.  
 XX WO200135984-A1.  
 XX PD 25-MAY-2001.  
 XX PF 17-NOV-2000; 2000WO-JP008119.  
 XX PR 18-NOV-1999; 99JP-00327900.  
 XX PR 26-SEP-2000; 2000JP-00297073.  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX Kitada C, Matsumoto H, Hinuma S;  
 PI  
 XX

DR WPI; 2001-355552/37.  
 XX Use of G protein receptor ligand or peptide for controlling corticotropin  
 PT releasing hormone secretion.  
 XX  
 PS Disclosure; Page 75; 90pp; Japanese.  
 XX  
 CC The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
 CC present sequence is a peptide used in the exemplification of the  
 CC invention  
 XX  
 SQ Sequence 21 AA;  
 Query Match 87.8%; Score 72; DB 4; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRPVGR 14  
 |||||  
 Db 7 AWYASRGIRPVGR 19  
 |||||  
 RESULT 16  
 ABU60847  
 ID ABU60847 standard; peptide; 21 AA.  
 XX  
 AC ABU60847;  
 XX  
 DT 06-MAY-2003 (first entry)  
 XX  
 DE Peptide production by gene recombination associated peptide #31.  
 XX  
 DE Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;  
 KW Gene recombination.  
 KW  
 OS Homo sapiens.  
 XX  
 PN WO200292829-A1.  
 XX  
 PD 21-NOV-2002.  
 XX  
 PF 16-MAY-2002; 2002WO-JP004735.  
 XX  
 PR 17-MAY-2001; 2001JP-00147341.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Nishimura O, Suenaga M, Ito T, Kitada C;  
 XX  
 DR WPI; 2003-129302/12.  
 XX  
 DR Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for  
 PT subsequent applications by gene recombination technique through tandem  
 PT repeats to provide precursor protein with specific cleavage sites.  
 XX  
 PS Disclosure; Page 69; 87pp; Japanese.  
 XX  
 CC The invention describes a method of producing a peptide comprising the  
 CC excision of the N and C-terminals of a target peptide with enzymes or  
 CC chemically through the attached cleavage sites repeated by ligation in a  
 CC precursor protein. The method is for producing (low-molecular) peptides  
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the  
 CC gene recombination technique through tandem repeats to provide a  
 CC precursor protein with specific cleavage sites. With this method, peptide  
 CC production can be carried out easily to provide large quantities of the  
 CC required peptides. This is the amino acid sequence of a peptide

CC associated with the peptide production method of the invention  
 XX  
 SQ Sequence 21 AA;  
 Query Match 87.8%; Score 72; DB 6; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRPVGR 14  
 |||||  
 Db 7 AWYASRGIRPVGR 19  
 |||||  
 RESULT 17  
 AAW31396  
 ID AAW31396 standard; peptide; 22 AA.  
 XX  
 AC AAW31396;  
 XX  
 DT 06-APR-1998 (first entry)  
 XX  
 DE Human type G protein-coupled receptor ligand fragment 6.  
 XX  
 DE G protein-coupled receptor; ligand binding; pharmaceutical; modulator;  
 KW pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9724436-A2.  
 XX  
 PD 10-JUL-1997.  
 XX  
 PF 26-DEC-1996; 96WO-JP003821.  
 XX  
 PR 28-DEC-1995; 95JP-00343371.  
 XX  
 PR 15-MAR-1996; 96JP-00059419.  
 XX  
 PR 12-AUG-1996; 96JP-00211805.  
 XX  
 PR 18-SEP-1996; 96JP-00246573.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;  
 PI Kitada C;  
 XX  
 DR WPI; 1997-363672/33.  
 XX  
 DR N-PSDB; AAV02433.  
 XX  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland.  
 XX  
 PS Claim 2; Page 186; 259pp; English.  
 XX  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the  
 CC sequence represented in AAW31396 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a prophylactic  
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,  
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,  
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,  
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 CC ankylosing lateral sclerosis, acute myocardial infarction, infertility,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the  
 CC ligand affecting activation of the G protein-coupled receptor protein  
 XX  
 SQ Sequence 22 AA;



Query Match 87.8%; Score 72; DB 2; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
 Db 7 AWYASRGIRPVGR 19

RESULT 18  
 AAB10367  
 ID AAB10367 standard; peptide; 22 AA.  
 AC AAB10367;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Human oxytocin secretion promoting peptide SEQ ID NO: 37.  
 XX  
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200038704-A1.  
 XX  
 PD 06-JUL-2000.  
 XX  
 XX 22-DEC-1999; 99WO-JP007199.  
 PF  
 XX 25-DEC-1998; 98JP-00369585.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Matsumoto H, Kitada C, Hinuma S;  
 PI  
 XX WPI; 2000-452298/39.  
 DR  
 XX Physiologically-active polypeptide recognized as ligand by G protein-  
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs  
 PT for diseases relating to oxytocin secretion and in veterinary medicine.  
 XX  
 PS Disclosure; Page 64; 72pp; Japanese.  
 XX  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak  
 CC pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter  
 XX  
 SQ Sequence 22 AA;

Query Match 87.8%; Score 72; DB 3; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
 Db 7 AWYASRGIRPVGR 19

RESULT 19  
 AAG62536  
 ID AAG62536 standard; peptide; 22 AA.  
 XX

AC AAG62536;  
 XX  
 DT 24-AUG-2001 (first entry)  
 XX  
 DE Human CRH releasing protein related peptide SEQ ID NO: 37.  
 XX  
 KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
 KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
 KW Addison's disease; adrenal gland hyperfunction; obesity.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200135984-A1.  
 XX  
 PD 25-MAY-2001.  
 XX  
 XX 17-NOV-2000; 2000WO-JP008119.  
 PF  
 XX 18-NOV-1999; 99JP-00327900.  
 PR  
 XX 26-SEP-2000; 2000JP-00297073.  
 PA  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Kitada C, Matsumoto H, Hinuma S;  
 XX  
 DR WPI; 2001-355552/37.  
 XX  
 XX Use of G protein receptor ligand or peptide for controlling corticotropin  
 PT releasing hormone secretion.  
 PT  
 XX Disclosure; Page 75; 90pp; Japanese.  
 XX  
 CC The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism;  
 CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
 CC present sequence is a peptide used in the exemplification of the  
 CC invention  
 XX  
 XX Sequence 22 AA;

Query Match 87.8%; Score 72; DB 4; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
 Db 7 AWYASRGIRPVGR 19

RESULT 20  
 ABU60848  
 ID ABU60848 standard; peptide; 22 AA.  
 XX  
 AC ABU60848;  
 XX  
 DT 06-MAY-2003 (first entry)  
 XX  
 DE Peptide production by gene recombination associated peptide #32.  
 XX  
 KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;  
 KW gene recombination.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200292829-A1.  
 XX  
 PD 21-NOV-2002.  
 XX

PF 16-MAY-2002; 2002WO-JP004735.  
 XX  
 PR 17-MAY-2001; 2001JP-00147341.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Nishimura O, Suenaga M, Ito T, Kitada C;  
 XX  
 XX WPI; 2003-129302/12.  
 DR  
 XX  
 XX Process for producing peptides e.g. Kiss-1 peptide and GPR8 ligand for  
 PT subsequent applications by gene recombination technique through tandem  
 PT repeats to provide precursor protein with specific cleavage sites.  
 XX  
 XX Disclosure; Page 69; 87pp; Japanese.  
 PS  
 XX The invention describes a method of producing a peptide comprising the  
 CC excision of the N and C-terminals of a target peptide with enzymes or  
 CC chemically through the attached cleavage sites repeated by ligation in a  
 CC precursor protein. The attached is for producing (low-molecular) peptides  
 CC e.g. Kiss-1 peptide and GPR8 ligand for subsequent applications by the  
 CC gene recombination technique through tandem repeats to provide a  
 CC precursor protein with specific cleavage sites. With this method, peptide  
 CC production can be carried out easily to provide large quantities of the  
 CC required peptides. This is the amino acid sequence of a peptide  
 CC associated with the peptide production method of the invention  
 XX  
 XX Sequence 22 AA;  
 SQ  
 Query Match 87.8%; Score 72; DB 6; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 AWYASRGIRPVGR 14  
 Db |||||  
 7 AWYASRGIRPVGR 19  
 |||||  
 RESULT 21  
 AAY49299  
 ID AAY49299 standard; peptide; 30 AA.  
 XX  
 AC AAY49299;  
 XX  
 XX 22-FEB-2000 (first entry)  
 DT  
 XX 19P2 ligand peptide fragment.  
 DE  
 XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
 KW  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 30  
 ET /note= "C-terminal amide"  
 FT  
 XX WO9960112-A1.  
 PN  
 XX 25-NOV-1999.  
 PD  
 XX 20-MAY-1999; 99WO-JP002650.  
 PF  
 XX 21-MAY-1998; 98JP-00140293.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Matsumoto H, Kitada C, Hinuma S;  
 PI  
 XX WPI; 2000-039381/03.  
 DR  
 XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying  
 XX diseases related to ligand abnormality.  
 PT

XX Disclosure; Page 27; 73pp; Japanese.  
 PS  
 XX The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in  
 CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand  
 XX  
 XX Sequence 30 AA;  
 SQ  
 Query Match 87.8%; Score 72; DB 3; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 AWYASRGIRPVGR 14  
 Db |||||  
 18 AWYASRGIRPVGR 30  
 |||||  
 RESULT 22  
 AAW31391  
 ID AAW31391 standard; peptide; 31 AA.  
 XX  
 AC AAW31391;  
 XX  
 XX 06-APR-1998 (first entry)  
 DT  
 XX Human type G protein-coupled receptor ligand fragment 1.  
 DE  
 XX G protein-coupled receptor; ligand binding; pharmaceutical; modulator;  
 KW pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9724436-A2.  
 PN  
 XX 10-JUL-1997.  
 DD  
 XX 26-DEC-1996; 96WO-JP003821.  
 PF  
 XX 28-DEC-1995; 95JP-00343371.  
 PR  
 XX 15-MAR-1996; 96JP-00059419.  
 PR  
 XX 12-AUG-1996; 96JP-00211805.  
 PR  
 XX 18-SEP-1996; 96JP-00246573.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;  
 PI Kitada C;  
 PI  
 XX WPI; 1997-363672/33.  
 DR N-PSDB; AAV02428.  
 XX  
 XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland.  
 PT  
 XX Claim 2; Page 184; 258pp; English.  
 PS  
 XX This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 53 of the  
 CC sequence represented in AAW31390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a prophylactic  
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC

CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,  
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,  
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,  
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 CC myelotrophic lateral sclerosis, acute myocardial infarction, infertility,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the  
 CC ligand affecting activation of the G protein-coupled receptor protein

XX  
 SQ Sequence 31 AA;

Query Match 87.8%; Score 72; DB 2; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
 |||||  
 Db 18 AWYASRGIRPVGR 30

RESULT 23

AAW87615  
 ID AAW87615 standard; peptide; 31 AA.

XX AC AAW87615;

XX DT 29-MAR-1999 (first entry)

XX DE Human 19P2 ligand.

XX KW 19P2 ligand; G protein coupled receptor; pituitary;  
 XX KW prolactin releasing peptide; human; dementia; breast cancer; therapy.

XX OS Homo sapiens.

XX FN EP887417-A2.

XX PD 30-DEC-1998.

XX PF 25-JUN-1998; 98EP-00111725.

XX PR 27-JUN-1997; 97JP-00172118.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Suenaga M, Moriya T, Tanaka Y, Nishimura O;

XX DR WPI; 1999-047884/05.

XX PT Producing a 19P2 pituitary G protein receptor ligand - by cleavage of a  
 XX PT fusion protein, useful for preventing and treating dementia, breast  
 XX PT cancer, renal failure and autoimmune disease.

XX PS Claim 5; Page 35; 56pp; English.

XX CC This is the amino acid sequence of the human pituitary G protein-coupled  
 CC receptor ligand 19P2L. A method suitable for commercial high-level  
 CC production of 19P2L comprises expressing the ligand in host cells as a  
 CC recombinant fusion protein e.g. with human basic fibroblast growth factor  
 CC (see AAV83796-97) that has been modified to include an N-terminal  
 CC cysteine residue. The ligand is released from the fusion by cyanylation  
 CC followed by ammonolysis. 19P2L has prolactin secretion-stimulating and  
 CC (at high doses) prolactin secretion-inhibiting properties. It can be used  
 CC in the treatment and prevention of various diseases including: senile  
 CC dementia, cerebrovascular dementia, and dementia associated with:  
 CC neurological disorders (e.g. Alzheimer's disease, Parkinson's disease,  
 CC Pick's disease, Huntington's disease), infectious diseases (e.g.  
 CC Creutzfeldt-Jakob's), endocrine or metabolic disease or toxicosis (e.g.  
 CC hypothyroidism, vitamin B12 deficiency, alcoholism, intoxication by  
 CC drugs, metal and organic compounds), tumorigenic diseases (e.g. brain  
 CC tumour), traumatic diseases (e.g. chronic subarachnoid haemorrhage, and  
 CC other types of dementia, depression, hyperactive child syndrome

CC (microencephalopathy) and disturbance of consciousness. It is also useful  
 CC for prevention and treatment of diseases associated with prolactin hypo  
 CC and hypersecretion respectively, including: hyperprolactinaemia,  
 CC pituitary adenoma, breast cancer, infertility, impotence and autoimmune  
 CC disease (hypersecretion disorders), and seminal vesicle hypoplasia,  
 CC osteoporosis, menopausal syndrome and renal failure (hyposecretion  
 CC disorders). The 19P2 polypeptide/amide is also useful as a test reagent  
 CC for study of the prolactin secretory function or as a lactagogue in  
 CC mammalian farm animals

XX SQ Sequence 31 AA;

Query Match 87.8%; Score 72; DB 2; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
 |||||  
 Db 18 AWYASRGIRPVGR 30

RESULT 24

AAW97235

ID AAW97235 standard; peptide; 31 AA.

XX AC AAW97235;

XX DT 06-MAY-1999 (first entry)

XX DE Human type ligand polypeptide fragment.

XX KW Rat type ligand; modulation; prolactin secretion;  
 XX KW G protein-coupled receptor; GPCR; hypovarianism; gonecyst cacogenesis;  
 XX KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
 XX KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
 XX KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 XX KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 XX KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 XX KW contraceptive; placental function; choriocarcinoma; hydatid mole;  
 XX KW irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 XX KW abnormal lipidmetabolism; oxytocia.

XX OS Homo sapiens.

XX FN WO9858962-A1.

XX PD 30-DEC-1998.

XX PF 22-JUN-1998; 98WO-JP002765.

XX PR 23-JUN-1997; 97JP-00165437.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Hinuma S, Kawamata Y, Fujii R, Matsumoto H;

XX DR WPI; 1999-105614/09.

XX PT Use of G protein-coupled receptor ligands - for modulating prolactin  
 XX PT secretion or placental function, e.g. for treating menopausal syndrome,  
 XX PT tumours, autoimmune disease or abnormal pregnancy.

XX PS Claim 3; Page 159; 241pp; English.

XX CC The present sequence represents a human type ligand fragment. It is used  
 CC in the course of the invention. The specification describes an agent for  
 CC modulating prolactin secretion which comprises a ligand polypeptide or a  
 CC salt, for a G protein-coupled receptor (GPCR) protein. The agents for  
 CC promoting prolactin secretion can be used for treating or preventing  
 CC hypovarianism, gonecyst cacogenesis, menopausal syndrome, euthyroid or  
 CC hypometabolism. They can be used for promoting lactation in a domestic  
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin  
 CC secretion can be used for treating or preventing pituitary adenomatosis,



DE XX Prolactin releasing peptide SEQ ID NO:165.

KW Protection; endogenous therapeutic peptide; peptidase; conjugation;

KW blood component; modification; succinimidyl; maleimido group; amino;

KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

OS Homo sapiens.

OS Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US013576.

XX 17-MAY-1999; 99US-0134406P.

XX 10-SEP-1999; 99US-0153406P.

XX 15-OCT-1999; 99US-0159783P.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents

XX peptidase degradation, useful for increasing length of in vivo activity.

XX Disclosure; Page 244; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I) comprising a therapeutically active amino acid region (III) and a reactive group (II) (e.g. succinimidyl and maleimido groups) attached to a less therapeutically active amino acid region (IV), which covalently bonds with amino/hydroxyl/thiol groups on blood components to form a peptidase stabilised therapeutic peptide composed of 3-50 amino acids. (I) are useful for modifying therapeutic peptides e.g. hormones, growth factors and neurotransmitters, to protect them from peptidase activity in vivo for the treatment of various disorders. Endogenous therapeutic peptides are not suitable as drug candidates as they require frequent administration due to rapid degradation by peptidases in the body. Modifying and attaching therapeutic peptides to albumin prevents or reduces the action of peptidases to increase length of activity (half life) and specificity as bonding to large molecules decreases intracellular uptake and interference with physiological processes. AAB90829 to AAB92441 represent peptides which can be used in the exemplification of the present invention

XX SQ Sequence 31 AA;

Query Match 87.8%; Score 72; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 2.4e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14

DB 18 AWYASRGIRPVGR 30

RESULT 29

AAB92531

ID AAG62531 standard; peptide; 31 AA.

XX AC

XX AAG62531;

XX 24-AUG-2001 (first entry)

XX Human CRH releasing protein related peptide SEQ ID NO: 32.

XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand; analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism; Addison's disease; adrenal gland hyperfunction; obesity.

XX Homo sapiens.

XX WO200135984-A1.

XX 25-MAY-2001.

DE XX Prolactin releasing peptide SEQ ID NO:169.

KW Protection; endogenous therapeutic peptide; peptidase; conjugation;

KW blood component; modification; succinimidyl; maleimido group; amino;

KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

OS Homo sapiens.

OS Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US013576.

XX 17-MAY-1999; 99US-0134406P.

XX 10-SEP-1999; 99US-0153406P.

XX 15-OCT-1999; 99US-0159783P.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents

XX peptidase degradation, useful for increasing length of in vivo activity.

XX Disclosure; Page 244; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I) comprising a therapeutically active amino acid region (III) and a reactive group (II) (e.g. succinimidyl and maleimido groups) attached to a less therapeutically active amino acid region (IV), which covalently bonds with amino/hydroxyl/thiol groups on blood components to form a peptidase stabilised therapeutic peptide composed of 3-50 amino acids. (I) are useful for modifying therapeutic peptides e.g. hormones, growth factors and neurotransmitters, to protect them from peptidase activity in vivo for the treatment of various disorders. Endogenous therapeutic peptides are not suitable as drug candidates as they require frequent administration due to rapid degradation by peptidases in the body. Modifying and attaching therapeutic peptides to albumin prevents or reduces the action of peptidases to increase length of activity (half life) and specificity as bonding to large molecules decreases intracellular uptake and interference with physiological processes. AAB90829 to AAB92441 represent peptides which can be used in the exemplification of the present invention

XX SQ Sequence 31 AA;

Query Match 87.8%; Score 72; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 2.4e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14

DB 18 AWYASRGIRPVGR 30

RESULT 28

AAB90995

ID AAB90995 standard; peptide; 31 AA.

XX AC

XX AAB90995;

XX 22-JUN-2001 (first entry)

XX Prolactin releasing peptide SEQ ID NO:169.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;

XX blood component; modification; succinimidyl; maleimido group; amino;

XX hydroxyl; thiol; hormone; growth factor; neurotransmitter.

PF 17-NOV-2000; 2000WO-JP008119.  
 XX  
 PR 18-NOV-1999; 99JP-00327900.  
 PR 26-SEP-2000; 2000JP-00297073.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Kitada C, Matsumoto H, Hinuma S;  
 XX  
 DR WPI; 2001-355552/37.  
 XX  
 XX  
 PT Use of G protein receptor ligand or peptide for controlling corticotropin  
 PT releasing hormone secretion.  
 XX  
 PS Claim 3; Page 73-74; 90pp; Japanese.  
 XX  
 CC The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
 CC present sequence is a peptide used in the exemplification of the  
 CC invention  
 CC  
 XX  
 SQ Sequence 31 AA;  
 Query Match 87.8%; Score 72; DB 4; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRPVGR 14  
 |||||  
 DB 18 AWYASRGIRPVGR 30  
 |||||  
 RESULT 30  
 AA26401  
 ID AAE26401 standard; peptide; 31 AA.  
 AC AAE26401;  
 XX  
 DT 13-DEC-2002 (first entry)  
 XX  
 DE Human PrRP-31 peptide.  
 XX  
 KW Human; wakefulness; sleep disorder; prolactin releasing peptide receptor;  
 KW PrRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;  
 KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;  
 KW anticonvulsant.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002037533-A1.  
 XX  
 PD 28-MAR-2002.  
 XX  
 PF 17-AUG-2001; 2001US-00932161.  
 XX  
 PR 28-APR-2000; 2000US-00560915.  
 XX  
 PA (CIVE/) CIVELLI O.  
 PA (LINS/) LIN S.  
 XX  
 PI Civelli O, Lin S;  
 XX  
 DR WPI; 2002-403931/43.  
 XX  
 PT Screening for compounds useful for promoting wakefulness or sleep, and  
 PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep  
 PT apnea, comprises administering a prolactin releasing peptide agonist or

PT antagonist.  
 XX  
 PS Disclosure; Page 24; 35pp; English.  
 XX  
 CC The present invention relates to a method of screening for compounds for  
 CC promoting wakefulness or sleep in a mammal. The method involves  
 CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)  
 CC agonist or antagonist respectively and determining the ability of the  
 CC compound to promote wakefulness or sleep. The compounds identified from  
 CC the method are used in the therapy of epilepsy and other diseases  
 CC associated with absence seizures and in promoting wakefulness and sleep  
 CC in individuals having sleep disorders such as insomnia and narcolepsy.  
 CC PrRP receptor agonists may be used to treat common disorders which lead  
 CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and  
 CC psychogenic hypersomnia. PrRP receptor antagonists are useful for  
 CC promoting sleep and for treating insomnia such as adjustment sleep  
 CC disorder and psychophysiological insomnia. The present sequence is human  
 CC PrRP-31 peptide  
 XX  
 SQ Sequence 31 AA;  
 Query Match 87.8%; Score 72; DB 5; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRPVGR 14  
 |||||  
 DB 18 AWYASRGIRPVGR 30  
 |||||  
 RESULT 31  
 ABU60843  
 ID ABU60843 standard; peptide; 31 AA.  
 AC ABU60843;  
 XX  
 XX  
 DT 06-MAY-2003 (first entry)  
 XX  
 DE Peptide production by gene recombination associated peptide #27.  
 XX  
 KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;  
 KW gene recombination.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200292829-A1.  
 XX  
 PD 21-NOV-2002.  
 XX  
 PF 16-MAY-2002; 2002WO-JP004735.  
 XX  
 PR 17-MAY-2001; 2001JP-00147341.  
 XX  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 XX Nishimura O, Suenaga M, Ito T, Kitada C;  
 PI  
 DR WPI; 2003-129302/12.  
 XX  
 XX  
 PT Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for  
 PT subsequent applications by gene recombination technique through tandem  
 PT repeats to provide precursor protein with specific cleavage sites.  
 XX  
 PS Disclosure; Page 67; 87pp; Japanese.  
 XX  
 CC The invention describes a method of producing a peptide comprising the  
 CC excision of the N and C-terminals of a target peptide with enzymes or  
 CC chemically through the attached cleavage sites repeated by ligation in a  
 CC precursor protein. The method is for producing (low-molecular) peptides  
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the  
 CC gene recombination technique through tandem repeats to provide a  
 CC precursor protein with specific cleavage sites. With this method, peptide  
 CC production can be carried out easily to provide large quantities of the

CC required peptides. This is the amino acid sequence of a peptide  
 CC associated with the peptide production method of the invention

SQ Sequence 31 AA;

Query Match 87.8%; Score 72; DB 6; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
 |||||  
 DB 18 AWYASRGIRPVGR 30

RESULT 32

ABU60827  
 ID ABU60827 standard; peptide; 31 AA.

AC ABU60827;

DT 06-MAY-2003 (first entry)

DE Peptide production by gene recombination associated peptide #11.  
 KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;  
 KW Gene recombination.

OS Homo sapiens.

PN WO200292829-A1.

XX 21-NOV-2002.

PF 16-MAY-2002; 2002WO-JP004735.

XX 17-MAY-2001; 2001JP-00147341.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Nishimura O, Suenaga M, Ito T, Kitada C;

XX WPI; 2003-129302/12.

XX Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for  
 PT subsequent applications by gene recombination technique through tandem  
 PT repeats to provide precursor protein with specific cleavage sites.

XX Disclosure; Page 59; 87pp; Japanese.

XX The invention describes a method of producing a peptide comprising the  
 CC excision of the N and C-terminals of a target peptide with enzymes or  
 CC chemically through the attached cleavage sites repeated by ligation in a  
 CC precursor protein. The method is for producing (low-molecular) peptides  
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the  
 CC gene recombination technique through tandem repeats to provide a  
 CC precursor protein with specific cleavage sites. With this method, peptide  
 CC production can be carried out easily to provide large quantities of the  
 CC required peptides. This is the amino acid sequence of a peptide  
 CC associated with the peptide production method of the invention

XX Sequence 31 AA;

Query Match 87.8%; Score 72; DB 6; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14

DB 18 AWYASRGIRPVGR 30

RESULT 33

ADC71228

ID ADC71228 standard; peptide; 31 AA.

XX ADC71228;

XX 18-DEC-2003 (first entry)

XX Human peptide sequence 2 related to the human serine protease.

DE human; serine protease precursor; hormone; neurohypophyseal; diabetes;  
 KW diabetic retinopathy; cataract; antidiabetic; ophthalmological.

XX Homo sapiens.

OS WO2003062429-A1.

PN 31-JUL-2003.

XX 22-JAN-2003; 2003WO-JP000547.

XX 23-JAN-2002; 2002JP-00013849.

XX 10-OCT-2002; 2002JP-00298003.

XX (YAMA ) YAMANOUCHI PHARM CO LTD.

XX Kagoshima M, Yamaji N, Takeda M, Abe K, Kawabe T;

XX WPI; 2003-598754/56.

XX Protease precursor for developing treatment for diabetes and cataracts.

XX Example 10; SEQ ID NO 13; 61pp; Japanese.

XX This invention relates to a novel human serine protease precursor  
 CC polypeptide. Specifically, it refers to a novel type II transmembrane  
 CC serine protease that participates in the control of hormones produced by  
 CC the pancreatic gland. As such, it can be used in a screening method for  
 CC the identification of compounds that are useful for the treatment of  
 CC diseases where a neurohypophyseal hormone participates, such as  
 CC diabetes, diabetic retinopathy and cataract. The present invention  
 CC describes the compounds of this invention as antidiabetic and  
 CC ophthalmological. This peptide sequence is human peptide 2 related to the  
 CC human serine protease of the invention.

XX Sequence 31 AA;

Query Match 87.8%; Score 72; DB 7; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14

DB 18 AWYASRGIRPVGR 30

RESULT 34

AWW31392

ID AAW31392 standard; peptide; 32 AA.

XX AAW31392;

XX 06-APR-1998 (first entry)

XX Human type G protein-coupled receptor ligand fragment 2.

XX G protein-coupled receptor; ligand binding; pharmaceutical; modulator;  
 KW pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent.

OS Homo sapiens.

XX WO9724436-A2.

XX 10-JUL-1997.

XX PF 26-DEC-1996; 96WO-JF003821.  
 XX XX  
 PR 28-DEC-1995; 95JP-00343371.  
 PR 15-MAR-1996; 96JP-00059419.  
 PR 12-AUG-1996; 96JP-00211805.  
 PR 18-SEP-1996; 96JP-00246573.  
 XX XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;  
 PI Kitada C;  
 XX  
 DR WPI; 1997-363672/33.  
 DR N-PSDB; AAV02429.  
 XX  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland.  
 XX  
 PS Claim 2; Page 185; 258pp; English.  
 XX  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 54 of the  
 CC sequence represented in AAW1390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a prophylactic  
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,  
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,  
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,  
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the  
 CC ligand affecting activation of the G protein-coupled receptor protein  
 XX  
 XX SQ Sequence 32 AA;  
 Query Match 87.8%; Score 72; DB 2; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-05; Mismatches 0; Gaps 0;  
 Matches 13; Conservative 0; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRPVGR 14  
 |||||  
 Db 18 AWVASRGIRPVGR 30  
 |||||  
 RESULT 35  
 AAB10363  
 ID AAB10363 standard; peptide; 32 AA.  
 XX  
 AC AAB10363;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Human oxytocin secretion promoting peptide SEQ ID NO: 33.  
 XX  
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2000038704-A1.  
 XX  
 PD 06-JUL-2000.  
 XX  
 PF 22-DEC-1999; 99WO-JF007199.

XX PR 25-DEC-1998; 98JP-00369585.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Matsumoto H, Kitada C, Hinuma S;  
 XX  
 DR WPI; 2000-452298/39.  
 XX  
 PT Physiologically-active polypeptide recognized as ligand by G protein-  
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs  
 PT for diseases relating to oxytocin secretion and in veterinary medicine.  
 XX  
 PS Disclosure; Page 62; 72pp; Japanese.  
 XX  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak  
 CC pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter  
 XX  
 SQ Sequence 32 AA;  
 Query Match 87.8%; Score 72; DB 3; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-05; Mismatches 0; Gaps 0;  
 Matches 13; Conservative 0; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRPVGR 14  
 |||||  
 Db 18 AWVASRGIRPVGR 30  
 |||||  
 RESULT 36  
 AAG62532  
 ID AAG62532 standard; peptide; 32 AA.  
 XX  
 AC AAG62532;  
 XX  
 DT 24-AUG-2001 (first entry)  
 XX  
 DE Human CRH releasing protein related peptide SEQ ID NO: 33.  
 XX  
 KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
 KW Addison's disease; adrenal gland hyperfunction; obesity.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200135984-A1.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF 17-NOV-2000; 2000WO-JP008119.  
 XX  
 PR 18-NOV-1999; 99JP-00327900.  
 PR 26-SEP-2000; 2000JP-00297073.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Kitada C, Matsumoto H, Hinuma S;  
 XX  
 DR WPI; 2001-355552/37.  
 XX  
 PT Use of G protein receptor ligand or peptide for controlling corticotropin  
 PT releasing hormone secretion.  
 XX  
 PS Disclosure; Page 74; 90pp; Japanese.  
 XX



CC The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
 CC present sequence is a peptide used in the exemplification of the  
 CC invention  
 XX  
 XX Sequence 32 AA;

XX Query Match

XX Best Local Similarity 87.8%; Score 72; DB 4; Length 32;

XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14

Db 18 AWYASRGIRPVGR 30

RESULT 37

ABU60844  
 ID ABU60844 standard; peptide; 32 AA.

XX AC ABU60844;

XX DT 06-MAY-2003 (first entry)

XX DE Peptide production by gene recombination associated peptide #28.

XX KW Peptide production; low-molecular peptide; KISS-1; GPR8 ligand;

XX Gene recombination.

XX OS Homo sapiens.

XX FN WO200292829-A1.

XX PD 21-NOV-2002.

XX PF 16-MAY-2002; 2002WO-JP004735.

XX PR 17-MAY-2001; 2001JP-00147341.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Nishimura O, Suenaga M, Ito T, Kitada C;

XX DR WPI; 2003-129302/12.

XX PT Process for producing peptides e.g. KISS-1 peptide and GPR8 ligand for  
 PT subsequent applications by gene recombination technique through tandem  
 PT repeats to provide precursor protein with specific cleavage sites.

XX PS Disclosure; Page 67; 87pp; Japanese.

XX The invention describes a method of producing a peptide comprising the  
 CC excision of the N and C-terminals of a target peptide with enzymes or  
 CC chemically through the attached cleavage sites repeated by ligation in a  
 CC precursor protein. The method is for producing (low-molecular) peptides  
 CC e.g. KISS-1 peptide and GPR8 ligand for subsequent applications by the  
 CC gene recombination technique through tandem repeats to provide a  
 CC precursor protein with specific cleavage sites. With this method, peptide  
 CC production can be carried out easily to provide large quantities of the  
 CC required peptides. This is the amino acid sequence of a peptide  
 CC associated with the peptide production method of the invention.

XX Sequence 32 AA;

XX Query Match

XX Best Local Similarity 87.8%; Score 72; DB 6; Length 32;

XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14

Db 18 AWYASRGIRPVGR 30

RESULT 38

AAW31393  
 ID AAW31393 standard; peptide; 33 AA.

XX AC AAW31393;

XX DT 06-APR-1998 (first entry)

XX DE Human type G protein-coupled receptor ligand fragment 3.

XX KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;  
 KW pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent.

XX OS Homo sapiens.

XX FN WO9724436-A2.

XX PD 10-JUL-1997.

XX PF 26-DEC-1996; 96WO-JP003821.

XX PR 28-DEC-1995; 95JP-00343371.

XX PR 15-MAR-1996; 96JP-00059419.

XX PR 12-AUG-1996; 96JP-00211805.

XX PR 18-SEP-1996; 96JP-00246573.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;

XX PI Kitada C;

XX DR WPI; 1997-363672/33.

XX DR N-PSDE; AAV02430.

XX PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland.

XX PS Claim 2; Page 185; 258pp; English.

XX This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 55 of the  
 CC sequence represented in AAW31390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a prophylactic  
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,  
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,  
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,  
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the  
 CC ligand affecting activation of the G protein-coupled receptor protein

XX Sequence 33 AA;

XX Query Match 87.8%; Score 72; DB 2; Length 33;

XX Best Local Similarity 100.0%; Pred. No. 2.5e-05;  
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14

|||||

Db 18 AWYASRGIRPVGR 30

RESULT 39  
AAB10364

ID AAB10364 standard; peptide; 33 AA.

XX AC AAB10364;

XX DT 24-NOV-2000 (first entry)

XX DE Human oxytocin secretion promoting peptide SEQ ID NO: 34.

XX KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;

XX KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;

XX KW caesarean section; artificial fertilization; galactostasis; goat; pig;

XX KW veterinary medicine; milk production.

XX OS Homo sapiens.

XX PN WO200038704-A1.

XX PD 06-JUL-2000.

XX PF 22-DEC-1999; 99WO-JP007199.

XX PR 25-DEC-1998; 98JP-00369585.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Matsumoto H, Kitada C, Hinuma S;

XX DR WPI; 2000-452298/39.

XX PT Physiologically-active polypeptide recognized as ligand by G protein-

XX PT coupled receptor protein, for promoting secretion of oxytocin, as drugs

XX PT for diseases relating to oxytocin secretion and in veterinary medicine.

XX PS Disclosure; Page 62-63; 72pp; Japanese.

XX CC This invention describes a novel oxytocin secretion-regulating agent

XX CC which contains a ligand peptide or its salt for the G protein-coupled

XX CC receptor protein. It is useful in the form of drugs for ameliorating,

XX CC preventing and treating diseases relating to oxytocin secretion e.g. weak

XX CC pains and atonic bleeding, before and after expulsion of placenta,

XX CC uterine recovery failure, caesarean section, stoppage of artificial

XX CC fertilization or galactostasis and is also applicable in veterinary

XX CC medicine for promoting milk production in cow, goat and pig. This

XX CC sequence represents a human peptide which acts as an oxytocin secretion

XX CC promoter

XX SQ Sequence 33 AA;

Query Match 87.8%; Score 72; DB 3; Length 33;  
Best Local Similarity 100.0%; Pred. No. 2.5e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | | | | | | | |  
Db 18 AWYASRGIRPVGR 30

RESULT 40  
AAG62533

ID AAG62533 standard; peptide; 33 AA.

XX AC AAG62533;

XX DT 24-AUG-2001 (first entry)

XX DE Human CRH releasing protein related peptide SEQ ID NO: 34.

XX KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;

KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.

XX OS Homo sapiens.

XX PN WO200135984-A1.

XX PD 25-MAY-2001.

XX PF 17-NOV-2000; 2000WO-JP008119.

XX PR 18-NOV-1999; 99JP-00327900.

XX PR 26-SEP-2000; 2000JP-00297073.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Kitada C, Matsumoto H, Hinuma S;

XX DR WPI; 2001-355552/37.

XX PT Use of G protein receptor ligand or peptide for controlling corticotropin

XX PT releasing hormone secretion.

XX PS Disclosure; Page 74; 90pp; Japanese.

XX CC The present sequence describes a method of controlling the secretion of

XX CC corticotrophin releasing hormone (CRH), involving the use of a G protein

XX CC receptor ligand. This can be used to control the secretion of CRH and is

XX CC useful as an analgesic or for treating, preventing or ameliorating

XX CC diseases associated with CRH secretion such as hyperaldosteronism,

XX CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's

XX CC disease (including boredom, nausea, pigmentation, hypogonadism, hair

XX CC loss, and hypotension), adrenal gland hypofunction and obesity. The

XX CC present sequence is a peptide used in the exemplification of the

XX CC invention

XX SQ Sequence 33 AA;

Query Match 87.8%; Score 72; DB 4; Length 33;  
Best Local Similarity 100.0%; Pred. No. 2.5e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | | | | | | | |  
Db 18 AWYASRGIRPVGR 30

Search completed: February 25, 2004, 06:42:35  
Job time : 55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 06:41:30 ; Search time 23 Seconds  
(without alignments)  
33.669 Million cell updates/sec

Title: US-09-700-643A-7

Perfect score: 82

Sequence: 1 CAWVASRGIRPVGRX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000 .

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep.\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	77	93.9	15	3	US-08-776-971-93
2	72	87.8	20	3	US-09-105-678A-46
3	72	87.8	20	3	US-08-776-971-64
4	72	87.8	20	3	US-09-421-208-46
5	72	87.8	20	4	US-09-560-915-18
6	72	87.8	21	3	US-09-105-678A-47
7	72	87.8	21	3	US-08-776-971-65
8	72	87.8	21	3	US-09-421-208-47
9	72	87.8	22	3	US-09-105-678A-48
10	72	87.8	22	3	US-08-776-971-66
11	72	87.8	22	3	US-09-421-208-48
12	72	87.8	31	3	US-09-105-678A-9
13	72	87.8	31	3	US-09-105-678A-43
14	72	87.8	31	3	US-08-776-971-61
15	72	87.8	31	3	US-09-421-208-9
16	72	87.8	31	4	US-09-421-208-43
17	72	87.8	31	4	US-09-560-915-15
18	72	87.8	32	3	US-09-105-678A-44
19	72	87.8	32	3	US-08-776-971-62
20	72	87.8	32	3	US-09-421-208-44
21	72	87.8	33	3	US-09-105-678A-45
22	72	87.8	33	3	US-08-776-971-63
23	72	87.8	33	3	US-09-421-208-45
24	72	87.8	87	3	US-08-776-971-59
25	72	87.8	87	3	US-08-776-971-135
26	72	87.8	87	3	US-08-776-971-138
27	68	82.9	19	3	US-09-105-678A-30

28	68	82.9	19	3	US-08-776-971-4	Sequence 4, Appli
29	68	82.9	19	3	US-09-421-208-30	Sequence 30, Appl
30	68	82.9	20	3	US-09-105-678A-34	Sequence 34, Appl
31	68	82.9	20	3	US-08-776-971-8	Sequence 8, Appli
32	68	82.9	20	3	US-08-776-971-98	Sequence 98, Appl
33	68	82.9	20	3	US-09-421-208-34	Sequence 34, Appl
34	68	82.9	20	4	US-09-560-915-16	Sequence 16, Appl
35	68	82.9	21	3	US-09-105-678A-35	Sequence 35, Appl
36	68	82.9	21	3	US-08-776-971-9	Sequence 9, Appli
37	68	82.9	21	3	US-09-421-208-35	Sequence 35, Appl
38	68	82.9	22	3	US-09-105-678A-36	Sequence 36, Appl
39	68	82.9	22	3	US-08-776-971-10	Sequence 10, Appl
40	68	82.9	22	3	US-09-421-208-36	Sequence 36, Appl
41	68	82.9	31	3	US-09-105-678A-7	Sequence 7, Appli
42	68	82.9	31	3	US-09-105-678A-31	Sequence 31, Appl
43	68	82.9	31	3	US-08-776-971-5	Sequence 5, Appli
44	68	82.9	31	3	US-08-776-971-97	Sequence 97, Appl
45	68	82.9	31	3	US-09-421-208-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-08-776-971-93

; Sequence 93, Application US/08776971B

; Patent No. 6228984

; GENERAL INFORMATION:

; APPLICANT: Hinuma, Shuji

; Habata, Yugo

; Kawanata, Yuji

; Hosoya, Masaki

; Fujii, Ryo

; Fukusumi, Shoji

; Kitada, Chieko

; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE

; NUMBER OF SEQUENCES: 140

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; STREET: 130 Water Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/776.971B

; FILING DATE: 06-Feb-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/JP96/03821

; FILING DATE: 28-DEC-1996

; APPLICATION NUMBER: JP 7/343371

; FILING DATE: 28-DEC-1995

; APPLICATION NUMBER: JP 8/59419

; FILING DATE: 15-MAR-1996

; APPLICATION NUMBER: JP 8/211805

; FILING DATE: 12-AUG-1996

; APPLICATION NUMBER: JP 8/246573

; FILING DATE: 18-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Conlin, David G.

; REGISTRATION NUMBER: 27,026

; REFERENCE/DOCKET NUMBER: 47176

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400

; TELEFAX: 617-523-6440

; INFORMATION FOR SEQ ID NO: 93:

; SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 93:

US-08-776-971-93

Query Match 93.9%; Score 77; DB 3; Length 15;  
Best Local Similarity 92.9%; Pred. No. 1.2e-06;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAWYASRGIRPVGR 14  
DB 1 CAWYAGRGIRPVGR 14

## RESULT 2

US-09-105-678A-46  
Sequence 46, Application US/09105678A

Patent No. 6103882

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato

APPLICANT: Moriya, Takeo

APPLICANT: Tanaka, Yoko

APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105,678A

FILING DATE: 26-JUN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-105-678A-46

Query Match 87.8%; Score 72; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
DB 7 AWYASRGIRPVGR 19

## RESULT 3

US-08-776-971-64  
Sequence 84, Application US/08776971B  
Patent No. 6228984

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji

APPLICANT: Habata, Yugo

APPLICANT: Kawanata, Yuji

APPLICANT: Hosoya, Masaki

APPLICANT: Fujii, Ryo

APPLICANT: Fukusumi, Shoji

APPLICANT: Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B

FILING DATE: 06-Feb-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821

FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371

FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419

FILING DATE: 15-MAR-1996

APPLICATION NUMBER: JP 8/211805

FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573

FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 47176

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 64:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 64:

US-08-776-971-84

Query Match 87.8%; Score 72; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
DB 7 AWYASRGIRPVGR 19

## RESULT 4

US-09-421-208-46

Sequence 46, Application US/09421208

Patent No. 6258561

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato

APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 87.8%; Score 72; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | | | | | | | |  
DB 7 AWYASRGIRPVGR 19

RESULT 5  
US-09-560-915-18  
; Sequence 18, Application US/09560915  
; Patent No. 6383764  
; GENERAL INFORMATION:  
; APPLICANT: Civeilli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Therapeutic Compositions and Methods  
; FILE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)  
; FILE REFERENCE: P-UC 3534  
; CURRENT APPLICATION NUMBER: US/09/560,915  
; CURRENT FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-09-560-915-18

Query Match 87.8%; Score 72; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | | | | | | | |  
DB 7 AWYASRGIRPVGR 19

RESULT 6  
US-09-105-678A-47  
; Sequence 47, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 87.8%; Score 72; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | | | | | | | |  
DB 7 AWYASRGIRPVGR 19

RESULT 7  
US-08-776-971-65  
; Sequence 65, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Habata, Yugo  
; APPLICANT: Kawamata, Yuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Kitada, Chieko  
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Collin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLSCULE TYPE: peptide
US-09-421-208-47

Query Match 87.8%; Score 72; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels

QY 2 AWYASRGIRPVGR 14
DB 7 AWYASRGIRPVGR 19
|||||

RESULT 9
US-09-105-678A-48
Sequence 48, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Collin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400

```

```
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-48

Query Match      87.8%; Score 72; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
Db      7 AWYASRGIRPVGR 19

RESULT 10
US-08-776-971-66
; Sequence 66, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
;             Habata, Yugo
;             Kawamata, Yuji
;             Hosoya, Masaki
;             Fujii, Ryo
;             Fukusumi, Shoji
;             Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 66:
US-08-776-971-66

Query Match      87.8%; Score 72; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
Db      7 AWYASRGIRPVGR 19

RESULT 11
US-09-421-208-48
; Sequence 48, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-48

Query Match      87.8%; Score 72; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
Db      7 AWYASRGIRPVGR 19

RESULT 12
US-09-105-678A-9
; Sequence 9, Application US/09105678A
```





```

;
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE: linear
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-08-776-971-61

Query Match      87.8%; Score 72; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      18 AWYASRGIRPVGR 30

RESULT 15
US-09-421-208-9
; Sequence 9, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-9

;
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-9

```

```

Query Match      87.8%; Score 72; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      18 AWYASRGIRPVGR 30

RESULT 16
US-09-421-208-43
; Sequence 43, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-43

Query Match      87.8%; Score 72; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      18 AWYASRGIRPVGR 30

RESULT 17
US-09-560-915-15
; Sequence 15, Application US/09560915
; Patent No. 6383764
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven

```

;; TITLE OF INVENTION: Therapeutic Compositions and Methods  
;; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)

;; FILE REFERENCE: P-UC 3534  
;; CURRENT APPLICATION NUMBER: US/09/560,915  
;; CURRENT FILING DATE: 2000-04-28  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: FASTSEQ for Windows Version 4.0  
;; SEQ ID NO 15  
;; LENGTH: 31  
;; TYPE: PrP  
;; ORGANISM: Homo Sapien  
US-09-560-915-15

Query Match 87.8%; Score 72; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14  
Db 18 AWYASRGIRPVGR 30  
|||||

## RESULT 18

US-09-105-678A-44  
; Sequence 44, Application US/09105678A  
; Patent No. 6103882

## GENERAL INFORMATION:

;; APPLICANT: Suenaga, Masato  
;; APPLICANT: Moriya, Takeo  
;; APPLICANT: Tanaka, Yoko  
;; APPLICANT: Nishimura, Osamu

;; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

;; NUMBER OF SEQUENCES: 52

;; CORRESPONDENCE ADDRESS: 52

;; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

;; STREET: 130 Water Street

;; CITY: Boston

;; STATE: MA

;; COUNTRY: USA

;; ZIP: 02109

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent In Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/105,678A

;; FILING DATE: 26-JUN-1998

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: JP 172118/1997

;; FILING DATE: 27-JUN-1997

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Conlin, David G.

;; REGISTRATION NUMBER: 27,026

;; REFERENCE/DOCKET NUMBER: 48466-342

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 617-523-3400

;; TELEFAX: 617-523-6440

;; INFORMATION FOR SEQ ID NO: 44:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 32 amino acids

;; TYPE: amino acid

;; STRANDEDNESS:

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

US-09-105-678A-44

Query Match 87.8%; Score 72; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14  
|||||

Db 18 AWYASRGIRPVGR 30

## RESULT 19

US-08-776-971-62  
; Sequence 62, Application US/08776971B  
; Patent No. 6228984

## GENERAL INFORMATION:

;; APPLICANT: Hinuma, Shuji  
;; Habata, Yugo  
;; Kawamata, Yuji  
;; Hosoya, Masaki  
;; Fujii, Ryo  
;; Fukusumi, Shoji  
;; Kitada, Chieko

;; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE

;; NUMBER OF SEQUENCES: 140

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

;; STREET: 130 Water Street

;; CITY: Boston

;; STATE: MA

;; COUNTRY: USA

;; ZIP: 02109

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Diskette

;; COMPUTER: IBM compatible

;; OPERATING SYSTEM: DOS

;; SOFTWARE: FASTSEQ for Windows Version 2.0

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/776,971B

;; FILING DATE: 06-Feb-1997

;; CLASSIFICATION: <Unknown>

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PCT/JP96/03821

;; FILING DATE: 28-DEC-1996

;; APPLICATION NUMBER: JP 7/343371

;; FILING DATE: 28-DEC-1995

;; APPLICATION NUMBER: JP 8/59419

;; FILING DATE: 15-MAR-1996

;; APPLICATION NUMBER: JP 8/211805

;; FILING DATE: 12-AUG-1996

;; APPLICATION NUMBER: JP 8/246573

;; FILING DATE: 18-SEP-1996

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Conlin, David G.

;; REGISTRATION NUMBER: 27,026

;; REFERENCE/DOCKET NUMBER: 47176

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 617-523-3400

;; TELEFAX: 617-523-6440

;; INFORMATION FOR SEQ ID NO: 62:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 32 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; FRAGMENT TYPE: internal

;; SEQUENCE DESCRIPTION: SEQ ID NO: 62:

Query Match 87.8%; Score 72; DB 3; Length 32;

Best Local Similarity 100.0%; Pred. No. 1.8e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14  
|||||

Db 18 AWYASRGIRPVGR 30  
|||||

RESULT 20,  
US-09-421-208-44

; Sequence 44, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-421-208-44  
  
Query Match 87.8%; Score 72; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | | | | | | | |  
DB 18 AWYASRGIRPVGR 30

RESULT 21  
US-09-105-678A-45  
; Sequence 45, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 33 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-45  
  
Query Match 87.8%; Score 72; DB 3; Length 33;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | | | | | | | |  
DB 18 AWYASRGIRPVGR 30

RESULT 22  
US-08-776-971-63  
; Sequence 63, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Fuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805

```
;
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Corlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-08-776-971-63

Query Match      87.8%; Score 72; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 AWYASRGIRPVGR 14
Db      18 AWYASRGIRPVGR 30

RESULT 23
US-09-421-208-45
; Sequence 45, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Corlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; QUERY MATCH      87.8%; Score 72; DB 3; Length 33;
; BEST LOCAL SIMILARITY 100.0%; Pred. No. 1.9e-05;
; MATCHES 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Corlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-08-776-971-63

Query Match      87.8%; Score 72; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 AWYASRGIRPVGR 14
Db      18 AWYASRGIRPVGR 30

RESULT 24
US-08-776-971-59
; Sequence 59, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Habata, Yugo
; APPLICANT: Kawamata, Yuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Fukusumi, Shoiji
; APPLICANT: Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Corlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-08-776-971-59

Query Match      87.8%; Score 72; DB 3; Length 87;
```

Best Local Similarity 100.0%; Pred. No. 5.2e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
|||||

Db 40 AWYASRGIRPVGR 52

## RESULT 25

US-08-776-971-135  
; Sequence 135, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 135:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 135:  
US-08-776-971-135

Query Match 87.8%; Score 72; DB 3; Length 87;  
Best Local Similarity 100.0%; Pred. No. 5.2e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
|||||

Db 40 AWYASRGIRPVGR 52

## RESULT 26

US-08-776-971-138  
; Sequence 138, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 138:  
US-08-776-971-138

Query Match 87.8%; Score 72; DB 3; Length 87;  
Best Local Similarity 100.0%; Pred. No. 5.2e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
|||||

Db 40 AWYASRGIRPVGR 52

## RESULT 27

US-09-105-678A-30

; Sequence 30, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-30

Query Match 82.9%; Score 68; DB 3; Length 19;  
Best Local Similarity 92.3%; Pred. No. 5e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14  
Db 7 AWYAGRGIRPVGR 19

RESULT 28  
US-08-776-971-4  
; Sequence 4, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Habata, Yugo  
; APPLICANT: Kawamata, Yuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Kitada, Chieko  
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-08-776-971-4

Query Match 82.9%; Score 68; DB 3; Length 19;  
Best Local Similarity 92.3%; Pred. No. 5e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14  
Db 7 AWYAGRGIRPVGR 19

RESULT 29  
US-09-421-208-30  
; Sequence 30, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:

US-09-105-678A-34

Query Match 82.9%; Score 68; DB 3; Length 20;  
Best Local Similarity 92.3%; Pred. No. 5.3e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ANYASRGIRPVGR 14  
Db 7 ANYAGRGIRPVGR 19

RESULT 31

US-08-776-971-8  
; Sequence 8, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-776-971-8

Query Match 82.9%; Score 68; DB 3; Length 20;  
Best Local Similarity 92.3%; Pred. No. 5.3e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

US-09-105-678A-34

Query Match 82.9%; Score 68; DB 3; Length 19;  
Best Local Similarity 92.3%; Pred. No. 5e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

QY 2 ANYASRGIRPVGR 14  
Db 7 ANYAGRGIRPVGR 19

RESULT 30

US-09-105-678A-34  
; Sequence 34, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Meriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | |  
Db 7 AWYAGRGIRPVGR 19

## RESULT 32

US-08-776-971-98  
; Sequence 98, Application US/08776971B  
; Patent No. 6228984

## GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA

COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776.971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/J996/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

## ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

## INFORMATION FOR SEQ ID NO: 98:

SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 98:

## US-08-776-971-98

Query Match 82.9%; Score 68; DB 3; Length 20;  
Best Local Similarity 92.3%; Pred. No. 5.3e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | |  
Db 7 AWYAGRGIRPVGR 19

## RESULT 33

## US-09-421-208-34

; Sequence 34, Application US/09421208  
; Patent No. 6258561

## GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA

COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

## INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

## US-09-421-208-34

Query Match 82.9%; Score 68; DB 3; Length 20;  
Best Local Similarity 92.3%; Pred. No. 5.3e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | |

Db 7 AWYAGRGIRPVGR 19  
| | | | | | | | | |

## RESULT 34

US-09-560-915-16  
; Sequence 16, Application US/09560915  
; Patent No. 6383764  
GENERAL INFORMATION:  
APPLICANT: Clivelli, Olivier  
; APPLICANT: Lin, Steven  
TITLE OF INVENTION: Therapeutic Compositions and Methods  
FILE REFERENCE: P-UC 3534  
CURRENT APPLICATION NUMBER: US/09/560,915  
CURRENT FILING DATE: 2000-04-28

## NUMBER OF SEQ ID NOS: 24

SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16  
LENGTH: 20  
TYPE: PPT  
ORGANISM: Bos taurus



US-09-560-915-16

Query Match 82.9%; Score 68; DB 4; Length 20;  
Best Local Similarity 92.3%; Pred. No. 5.3e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | |  
DB 7 AWYAGRGIRPVGR 19

RESULT 35

US-09-105-678A-35  
; Sequence 35, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-35

Query Match 82.9%; Score 68; DB 3; Length 21;  
Best Local Similarity 92.3%; Pred. No. 5.5e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | |  
DB 7 AWYAGRGIRPVGR 19

RESULT 36

US-08-776-971-9  
; Sequence 9, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji

Hosoya, Masaki  
Fuji, RyoFukusumi, Shoji  
Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 9:

US-08-776-971-9  
Query Match 82.9%; Score 68; DB 3; Length 21;  
Best Local Similarity 92.3%; Pred. No. 5.5e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | |  
DB 7 AWYAGRGIRPVGR 19

RESULT 37

US-09-421-208-35  
; Sequence 35, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
MOLECULE TYPE: peptide  
US-09-421-208-35

Query Match 82.9%; Score 68; DB 3; Length 21;  
Best Local Similarity 92.3%; Pred. No. 5.5e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0;  
Gaps 0;

OY 2 AWYASRGIRPVGR 14  
DB 7 AWYAGRGIRPVGR 19

RESULT 38  
US-09-105-678A-36  
Sequence 36, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-36

Query Match 82.9%; Score 68; DB 3; Length 22;  
Best Local Similarity 92.3%; Pred. No. 5.8e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0;  
Gaps 0;

OY 2 AWYASRGIRPVGR 14  
DB 7 AWYAGRGIRPVGR 19

RESULT 39  
US-08-776-971-10  
Sequence 10, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-08-776-971-10

Query Match 82.9%; Score 68; DB 3; Length 22;  
Best Local Similarity 92.3%; Pred. No. 5.8e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | | | | | | | | | | | |  
Db 7 AWYAGRGIRPVGR 19

RESULT 40  
US-09-421-208-36  
; Sequence 36, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: Peptide  
US-09-421-208-36

Query Match 82.9%; Score 68; DB 3; Length 22;  
Best Local Similarity 92.3%; Pred. No. 5.8e-05;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | | | | | | | | | | | |  
Db 7 AWYAGRGIRPVGR 19

Search completed: February 25, 2004, 06:45:00  
Job time : 23 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 06:43:56 ; Search time 34 Seconds  
(without alignments)  
93.156 Million cell updates/sec

Title: US-09-700-643A-7  
Perfect score: 82  
Sequence: 1 CAWVASRGIRPVGRX 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	87.8	20	9	US-09-932-161-18
2	72	87.8	20	14	US-10-044-592-41
3	72	87.8	31	9	US-09-932-161-15
4	72	87.8	31	14	US-10-044-592-41
5	72	87.8	17	13	US-10-044-592-92
6	68	82.9	19	13	US-10-044-592-27
7	68	82.9	20	9	US-09-932-161-16
8	68	82.9	20	13	US-10-044-592-42
9	68	82.9	20	14	US-10-096-777-16
10	68	82.9	21	13	US-10-044-592-43
11	68	82.9	22	13	US-10-044-592-44
12	68	82.9	31	9	US-09-932-161-13
13	68	82.9	31	13	US-10-044-592-39
14	68	82.9	31	14	US-10-096-777-13
15	68	82.9	32	13	US-10-044-592-40

16	68	82.9	33	13	US-10-044-592-41	Sequence 41, Appl
17	68	82.9	98	13	US-10-044-592-28	Sequence 28, Appl
18	68	82.9	98	13	US-10-044-592-38	Sequence 38, Appl
19	68	82.9	98	13	US-10-044-592-82	Sequence 82, Appl
20	68	82.9	98	13	US-10-044-592-84	Sequence 84, Appl
21	68	82.9	98	13	US-10-044-592-86	Sequence 86, Appl
22	68	82.9	98	13	US-10-044-592-88	Sequence 88, Appl
23	64	78.0	20	9	US-09-932-161-17	Sequence 17, Appl
24	64	78.0	20	13	US-10-044-592-6	Sequence 6, Appl
25	64	78.0	20	14	US-10-096-777-17	Sequence 17, Appl
26	64	78.0	31	9	US-09-932-161-14	Sequence 14, Appl
27	64	78.0	31	13	US-10-044-592-4	Sequence 4, Appl
28	64	78.0	31	13	US-10-044-592-5	Sequence 5, Appl
29	64	78.0	31	14	US-10-096-777-14	Sequence 14, Appl
30	64	78.0	70	13	US-10-044-592-90	Sequence 90, Appl
31	64	78.0	82	13	US-10-044-592-1	Sequence 1, Appl
32	64	78.0	86	13	US-10-044-592-96	Sequence 96, Appl
33	64	78.0	91	13	US-10-044-592-94	Sequence 94, Appl
34	63	76.8	29	13	US-10-044-592-26	Sequence 26, Appl
35	53	64.6	25	13	US-10-044-592-78	Sequence 78, Appl
36	46	56.1	647	15	US-10-093-463-188	Sequence 188, App
37	44	53.7	41	15	US-10-369-493-12939	Sequence 12939, A
38	44	53.7	555	15	US-10-369-493-8457	Sequence 8457, Ap
39	43	52.4	402	14	US-10-156-761-10748	Sequence 10748, A
40	43	52.4	592	14	US-10-156-761-14927	Sequence 14927, A
41	43	52.4	838	14	US-10-156-761-10342	Sequence 10342, A
42	42	51.2	10	13	US-10-044-592-9	Sequence 9, Appl
43	42	51.2	288	14	US-10-156-761-11812	Sequence 11812, A
44	42	51.2	645	10	US-09-934-455-120	Sequence 120, App
45	42	51.2	645	14	US-10-278-536-200	Sequence 200, App

ALIGNMENTS

RESULT 1  
US-09-932-161-18  
; Sequence 18 Application US/09932161  
; Patent No. US20020037533A1  
; GENERAL INFORMATION:  
; APPLICANT: Civeilli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Screening and Therapeutic Methods For Promoting Wakefulness and Sleep  
; FILE REFERENCE: P-UC 4679  
; CURRENT APPLICATION NUMBER: US/09/932,161  
; PRIOR FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: US 09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-09-932-161-18

Query Match 87.8% Score 72: DB 9: Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.2e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AWVASRGIRPVGR 14  
Db 7 AWVASRGIRPVGR 19

RESULT 2  
US-10-096-777-18  
; Sequence 18, Application US/10096777  
; Publication No. US20030171270A1  
; GENERAL INFORMATION:  
; APPLICANT: Civeilli, Olivier  
; APPLICANT: Lin, Steven

; TITLE OF INVENTION: Therapeutic Compositions and Methods  
; FILE REFERENCE: P-UC 3534

; CURRENT APPLICATION NUMBER: US/10/096,777

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: US/09/560,915

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: Fast-SEQ for Windows Version 4.0

; SEQ ID NO 18

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-096-777-18

Query Match 87.8%; Score 72; DB 14; Length 20;

Best Local Similarity 100.0%; Pred. No. 6.2e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14

|||||

Db 7 AWYASRGIRPVGR 19

#### RESULT 3

US-09-932-161-15

; Sequence 15, Application US/09932161

; Patent No. US20020037533A1

; GENERAL INFORMATION:

; APPLICANT: Civealli, Olivier

; APPLICANT: Lin, Steven

; TITLE OF INVENTION: Screening and Therapeutic Methods For

; FILE REFERENCE: P-UC 4679

; CURRENT APPLICATION NUMBER: US/09/932,161

; CURRENT FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: US 09/560,915

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: Fast-SEQ for Windows Version 4.0

; SEQ ID NO 15

; LENGTH: 31

; TYPE: PRT

; ORGANISM: Homo Sapien

US-09-932-161-15

Query Match 87.8%; Score 72; DB 9; Length 31;

Best Local Similarity 100.0%; Pred. No. 9.2e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14

|||||

Db 18 AWYASRGIRPVGR 30

#### RESULT 4

US-10-096-777-15

; Sequence 15, Application US/10096777

; Publication No. US20030171270A1

; GENERAL INFORMATION:

; APPLICANT: Civealli, Olivier

; APPLICANT: Lin, Steven

; TITLE OF INVENTION: Therapeutic Compositions and Methods

; FILE REFERENCE: P-UC 3534

; CURRENT APPLICATION NUMBER: US/10/096,777

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: US/09/560,915

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: Fast-SEQ for Windows Version 4.0

; SEQ ID NO 15

; LENGTH: 31

; TYPE: PRT  
; ORGANISM: Homo Sapien

US-10-096-777-15

Query Match 87.8%; Score 72; DB 14; Length 31;

Best Local Similarity 100.0%; Pred. No. 9.2e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14

|||||

Db 18 AWYASRGIRPVGR 30

#### RESULT 5

US-10-044-592-92

; Sequence 92, Application US/10044592

; Publication No. US20020143152A1

; GENERAL INFORMATION:

; APPLICANT: Hinuma, Shuji

; APPLICANT: Fukusumi, Shoji

; TITLE OF INVENTION: Polypeptides, their Production and Use

; FILE REFERENCE: 2463US2P

; CURRENT APPLICATION NUMBER: US/10/044,592

; CURRENT FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 09/403639

; PRIOR FILING DATE: 1999-25-10

; PRIOR APPLICATION NUMBER: PCT/JP98/01923

; PRIOR FILING DATE: 1998-04-27

; PRIOR APPLICATION NUMBER: JP 9-109974

; PRIOR FILING DATE: 1997-04-28

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE:

; SEQ ID NO 92

; LENGTH: 87

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-044-592-92

Query Match 87.8%; Score 72; DB 13; Length 87;

Best Local Similarity 100.0%; Pred. No. 0.00024;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14

|||||

Db 40 AWYASRGIRPVGR 52

#### RESULT 6

US-10-044-592-27

; Sequence 27, Application US/10044592

; Publication No. US20020143152A1

; GENERAL INFORMATION:

; APPLICANT: Hinuma, Shuji

; APPLICANT: Fukusumi, Shoji

; TITLE OF INVENTION: Polypeptides, their Production and Use

; FILE REFERENCE: 2463US2P

; CURRENT APPLICATION NUMBER: US/10/044,592

; CURRENT FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 09/403639

; PRIOR FILING DATE: 1999-25-10

; PRIOR APPLICATION NUMBER: PCT/JP98/01923

; PRIOR FILING DATE: 1998-04-27

; PRIOR APPLICATION NUMBER: JP 9-109974

; PRIOR FILING DATE: 1997-04-28

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE:

; SEQ ID NO 27

; LENGTH: 19

; TYPE: PRT

; ORGANISM: Bovine

US-10-044-592-27

Query Match 82.9%; Score 68; DB 13; Length 19;

Best Local Similarity 92.3%; Pred. No. 0.00026;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | |  
Db 7 AWYAGRGIRPVGR 19

RESULT 7  
US-09-932-161-16  
; Sequence 16, Application US/09932161  
; Patent No. US20020037533A1  
; GENERAL INFORMATION:  
; APPLICANT: Civeilli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Screening and Therapeutic Methods For  
; Promoting Wakefulness and Sleep  
; FILE REFERENCE: P-UC 4679  
; CURRENT APPLICATION NUMBER: US/09/932,161  
; PRIOR FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: US 09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-09-932-161-16

Query Match 82.9%; Score 68; DB 9; Length 20;  
Best Local Similarity 92.3%; Pred. No. 0.00028;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | |  
Db 7 AWYAGRGIRPVGR 19

RESULT 8  
US-10-044-592-42  
; Sequence 42, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Fukusumi, Shoji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 42  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-044-592-42

Query Match 82.9%; Score 68; DB 13; Length 20;  
Best Local Similarity 92.3%; Pred. No. 0.00028;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | |  
Db 7 AWYAGRGIRPVGR 19

RESULT 9  
US-10-096-777-16  
; Sequence 16, Application US/10096777  
; Publication No. US20030171270A1  
; GENERAL INFORMATION:  
; APPLICANT: Civeilli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Therapeutic Compositions and Methods  
; Relating To Prolactin Releasing Peptide (P-RP)  
; FILE REFERENCE: P-UC 3534  
; CURRENT APPLICATION NUMBER: US/10/096,777  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US/09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-10-096-777-16

Query Match 82.9%; Score 68; DB 14; Length 20;  
Best Local Similarity 92.3%; Pred. No. 0.00028;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | |  
Db 7 AWYAGRGIRPVGR 19

RESULT 10  
US-10-044-592-43  
; Sequence 43, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Fukusumi, Shoji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 43  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-044-592-43

Query Match 82.9%; Score 68; DB 13; Length 21;  
Best Local Similarity 92.3%; Pred. No. 0.00029;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
| | | | | | | | | |  
Db 7 AWYAGRGIRPVGR 19

RESULT 11  
US-10-044-592-44  
; Sequence 44, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Fukusumi, Shoji  
; TITLE OF INVENTION: Polypeptides, their Production and Use

```
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 44
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-44

Query Match      82.9%; Score 68; DB 13; Length 22;
Best Local Similarity 92.3%; Pred. No. 0.0003;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AWYASRGIRPVGR 14
Db      7 AWYAGRGIRPVGR 19

RESULT 12
US-09-932-161-13
; Sequence 13, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civeili, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-932-161-13

Query Match      82.9%; Score 68; DB 9; Length 31;
Best Local Similarity 92.3%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AWYASRGIRPVGR 14
Db      18 AWYAGRGIRPVGR 30

RESULT 13
US-10-044-592-39
; Sequence 39, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 40
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-40

Query Match      82.9%; Score 68; DB 13; Length 22;
Best Local Similarity 92.3%; Pred. No. 0.0003;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AWYASRGIRPVGR 14
Db      7 AWYAGRGIRPVGR 19

RESULT 14
US-10-096-777-13
; Sequence 13, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civeili, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRp)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-096-777-13

Query Match      82.9%; Score 68; DB 14; Length 31;
Best Local Similarity 92.3%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AWYASRGIRPVGR 14
Db      18 AWYAGRGIRPVGR 30

RESULT 15
US-10-044-592-40
; Sequence 40, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 40
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-40
```





```

RESULT 20
US-10-044-592-84
; Sequence 84, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; SOFTWARE:
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 84
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-84

Query Match      82.9%; Score 68; DB 13; Length 98;
Best Local Similarity 92.3%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      40 AWYAGRGIRPVGR 52

RESULT 21
US-10-044-592-86
; Sequence 86, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; SOFTWARE:
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 86
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-86

Query Match      82.9%; Score 68; DB 13; Length 98;
Best Local Similarity 92.3%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      40 AWYAGRGIRPVGR 52

RESULT 22
US-10-044-592-88
; Sequence 88, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; SOFTWARE:
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 88
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-88

Query Match      82.9%; Score 68; DB 13; Length 98;
Best Local Similarity 92.3%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      40 AWYAGRGIRPVGR 52

```

```

; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 88
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-88

Query Match      82.9%; Score 68; DB 13; Length 98;
Best Local Similarity 92.3%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      40 AWYAGRGIRPVGR 52

RESULT 23
US-09-932-161-17
; Sequence 17, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Rattus
US-09-932-161-17

Query Match      78.0%; Score 64; DB 9; Length 20;
Best Local Similarity 84.6%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      7 AWYTGIRPVGR 19

RESULT 24
US-10-044-592-6
; Sequence 6, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; SOFTWARE:
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 84
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-84

Query Match      82.9%; Score 68; DB 13; Length 98;
Best Local Similarity 92.3%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
DB      40 AWYAGRGIRPVGR 52

```

```
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-26
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(20)
; OTHER INFORMATION: antigen
US-10-044-592-6

Query Match      78.0%; Score 64; DB 13; Length 20;
Best Local Similarity 84.6%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
Db      7 AWYTGGRIRPVGR 19

RESULT 25
US-10-096-777-17
; Sequence 17, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Rattus
US-10-096-777-17

Query Match      78.0%; Score 64; DB 14; Length 20;
Best Local Similarity 84.6%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
Db      7 AWYTGGRIRPVGR 19

RESULT 26
US-09-932-161-14
; Sequence 14, Application US/09932161
; Patent No. US2002003753A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14

; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rattus
US-09-932-161-14

Query Match      78.0%; Score 64; DB 9; Length 31;
Best Local Similarity 84.8%; Pred. No. 0.0019;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
Db      18 AWYTGGRIRPVGR 30

RESULT 27
US-10-044-592-4
; Sequence 4, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 4
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-4

Query Match      78.0%; Score 64; DB 13; Length 31;
Best Local Similarity 84.6%; Pred. No. 0.0019;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
Db      18 AWYTGGRIRPVGR 30

RESULT 28
US-10-044-592-5
; Sequence 5, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 5
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
```

; LOCATION: (1)..(31)  
; OTHER INFORMATION: antigen  
US-10-044-592-5

Query Match 78.0%; Score 64; DB 13; Length 31;  
Best Local Similarity 84.6%; Pred. No. 0.0019;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
Db 18 AWYTGGRIRPVGR 30

## RESULT 29

US-10-096-777-14  
; Sequence 14, Application US/10096777  
; Publication No. US20030171270A1  
; GENERAL INFORMATION:  
; APPLICANT: Civel, Olivier  
; APPLICANT: Liu, Steven  
; TITLE OF INVENTION: Therapeutic Compositions and Methods  
; FILE REFERENCE: Relating To Prolactin Releasing Peptide (PrRP)  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US/10/096,777  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Rattus  
US-10-096-777-14

Query Match 78.0%; Score 64; DB 14; Length 31;  
Best Local Similarity 84.6%; Pred. No. 0.0019;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
Db 18 AWYTGGRIRPVGR 30

## RESULT 30

US-10-044-592-90  
; Sequence 90, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Fukusumi, Shoji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US/10/044,592  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 90  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Rat  
US-10-044-592-90

Query Match 78.0%; Score 64; DB 13; Length 70;  
Best Local Similarity 84.6%; Pred. No. 0.004;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
Db 39 AWYTGGRIRPVGR 51

## RESULT 31

US-10-044-592-1  
; Sequence 1, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Fukusumi, Shoji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 1  
; LENGTH: 82  
; TYPE: PRT  
; ORGANISM: Murine  
US-10-044-592-1

Query Match 78.0%; Score 64; DB 13; Length 82;  
Best Local Similarity 84.6%; Pred. No. 0.0046;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
Db 38 AWYTGGRIRPVGR 50

## RESULT 32

US-10-044-592-96  
; Sequence 96, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Fukusumi, Shoji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US/10/044,592  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 96  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: mammalian  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1224)..(1243)  
; OTHER INFORMATION: Bracket region depicted in FIG 39.  
US-10-044-592-96

Query Match 78.0%; Score 64; DB 13; Length 86;  
Best Local Similarity 84.6%; Pred. No. 0.0048;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14

```

Db      39 AWYTGGRGIRPVGR 51
      ||| |||||
Query Match      76.8%; Score 63; DB 13; Length 29;
Best Local Similarity 91.7%; Pred. No. 0.0026;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 33
US-10-044-592-94
; Sequence 94, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1998-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 94
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Mammalian
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(31)
; OTHER INFORMATION: primer
; NAME/KEY: misc.feature
; LOCATION: (925)-(955)
; OTHER INFORMATION: primer
; NAME/KEY: misc.feature
; LOCATION: (1)-(955)
; OTHER INFORMATION: insert fragment of pmGB3 as depicted in FIG 36 and 37.
US-10-044-592-94

Query Match      78.0%; Score 64; DB 13; Length 91;
Best Local Similarity 84.6%; Pred. No. 0.005;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 AWYASRGIRPVGR 14
      ||| |||||
Db      39 AWYTGGRGIRPVGR 51
      ||| |||||

RESULT 34
US-10-044-592-26
; Sequence 26, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1998-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 26
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Bovine
; OTHER INFORMATION: primer
US-10-044-592-26

Query Match      64.6%; Score 53; DB 13; Length 25;
Best Local Similarity 90.0%; Pred. No. 0.095;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AWYASRGIRP 11
      ||| |||||
Db      16 AWYAGRGIRP 25
      ||| |||||

RESULT 35
US-10-044-592-78
; Sequence 78, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 78
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Bovine
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(20)
; OTHER INFORMATION: primer
; NAME/KEY: misc.feature
; LOCATION: (52)-(76)
; OTHER INFORMATION: primer
US-10-044-592-78

Query Match      64.6%; Score 53; DB 13; Length 25;
Best Local Similarity 90.0%; Pred. No. 0.095;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AWYASRGIRP 11
      ||| |||||
Db      16 AWYAGRGIRP 25
      ||| |||||

RESULT 36
US-10-093-463-188
; Sequence 188, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glennnda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zethusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine

```

```
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: No. US20030208039A1el Antibodies that Bind to Antigenic Polypepti
; FILE REFERENCE: 21402-290A (Cura 590AT)
; CURRENT APPLICATION NUMBER: US/10/093,463
; CURRENT FILING DATE: 2002-06-24
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 188
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-093-463-188

Query Match 56.1%; Score 46; DB 15; Length 647;
Best Local Similarity 54.3%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CANYASRGIRPVGR 14
||| | | | | | | |
Db 611 CANYASRGIRPVGR 624

RESULT 37
US-10-369-493-12939
; Sequence 12939, Application US/10369493
; APPLICANT: OMURA, SATOSHI

; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12939
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(411)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-12939

Query Match 53.7%; Score 44; DB 15; Length 411;
Best Local Similarity 53.8%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14
||| | | | | | | |
Db 281 AWYASRGIRPVGR 293

RESULT 38
US-10-369-493-8457
; Sequence 8457, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8457
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Ralstonia metallidurans
US-10-369-493-8457

Query Match 53.7%; Score 44; DB 15; Length 555;
Best Local Similarity 61.5%; Pred. No. 49;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14
||| | | | | | | |
Db 59 AWYASRGIRPVGR 71

RESULT 39
US-10-156-761-10748
; Sequence 10748, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
```

; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 10748  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-10748

Query Match 52.4%; Score 43; DB 14; Length 402;  
Best Local Similarity 47.4%; Pred. No. 53;  
Matches 9; Conservative 1; Mismatches 1; Indels 8; Gaps 1;

QY 3 WYASRG-----IRPVG 13  
DB 91 WYAPRGQLSLRAAEIRPIG 109

RESULT 40  
US-10-156-761-14927  
; Sequence 14927, Application US/10156761  
; Publication NO. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 14927  
; LENGTH: 592  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-14927

Query Match 52.4%; Score 43; DB 14; Length 592;  
Best Local Similarity 72.7%; Pred. No. 76;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 AWYASRGIRPV 12  
DB 103 AWMAARGRPV 113

Search completed: February 25, 2004, 06:49:23  
Job time : 35 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 25, 2004, 06:40:45 ; Search time 21 seconds  
(without alignments)  
68.708 Million cell updates/sec

Title: US-09-700-643A-7

Perfect score: 82

Sequence: 1 CAWVASRGIRPVGRX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	78.0	83	2 JC7607	prolactin-releasin
2	44	53.7	302	2 B70631	hypothetical prote
3	44	53.7	719	2 S61046	ARF1 protein - yea
4	42	51.2	105	2 A97505	hypothetical prote
5	42	51.2	105	2 A92723	hypothetical prote
6	42	51.2	113	2 A12484	hypothetical prote
7	42	51.2	132	2 B82824	protein-export mem
8	42	51.2	250	2 G83400	hypothetical prote
9	42	51.2	269	2 A95148	licD2 protein (imp
10	42	51.2	269	2 G98015	hypothetical prote
11	41	50.0	256	2 E95936	hypothetical prote
12	41	50.0	276	2 D70817	hypothetical prote
13	41	50.0	546	2 A2368	hypothetical prote
14	41	50.0	966	2 A93878	trwC protein - Esc
15	40.5	49.4	313	2 B95351	VirB6 type IV secr
16	40	48.8	267	2 B70837	hypothetical prote
17	40	48.8	398	2 B75254	acetate kinase - D
18	40	48.8	580	2 S54589	ASP2 protein - yea
19	40	48.8	788	2 I64045	recombination prot
20	40	48.8	790	2 T47959	hypothetical prote
21	39	47.6	171	2 F75556	probable acetyltra
22	39	47.6	181	2 A82305	hypothetical prote
23	39	47.6	184	2 T35841	probable membrane
24	39	47.6	260	2 H82925	hypothetical prote
25	39	47.6	315	2 A97709	formamidopyrimidin
26	39	47.6	333	2 H82852	hydroxybenzoate oc
27	39	47.6	374	2 G70947	hypothetical prote
28	39	47.6	424	2 B38176	samb protein - Sal
29	39	47.6	424	2 A51034	UV protection prot

glycerol-3-phospha  
sn-glycerol-3-phos  
sn-glycerol-3-phos  
sulfide dehydrogen  
alcohol dehydrogen  
319K protein ndvB  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hydrolase, probabl  
cathapsin L - Para  
carboxypeptidase B  
hypothetical prote  
conserved hypothet  
carboxypeptidase B  
hypothetical prote

30 39 47.6 452 1 JNECGT  
31 39 47.6 452 2 E91019  
32 39 47.6 452 2 G85863  
33 39 47.6 459 2 T43558  
34 39 47.6 600 2 D84258  
35 39 47.6 2870 2 A35548  
36 38 46.3 121 2 H72507  
37 38 46.3 194 2 D87357  
38 38 46.3 238 2 H72646  
39 38 46.3 284 2 F95120  
40 38 46.3 294 2 S68784  
41 38 46.3 306 1 CPBOB  
42 38 46.3 330 2 T42947  
43 38 46.3 343 2 E95986  
44 38 46.3 416 1 A42332  
45 38 46.3 456 2 P90576

## ALIGNMENTS

### RESULT 1

JC7607

prolactin-releasing peptide - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001

C:Accession: JC7607

R;Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T.

Biochem. Biophys. Res. Commun. 281, 53-56, 2001

A>Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mnl

A:Reference number: JC7607; MUID:21092785; PMID:11178959

A:Contents: Spleen

A:Accession: JC7607

A:Molecule type: DNA

A:Residues: 1-83 <YAM>

A:Cross-references: DDBJ:AB040612; DBJ:AB040613

C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior pi

release, and stimulation of ACTH secretion from the pituitary.

C:Genetics:

A:Gene: PrRP

A:Introns: 33/1

Query Match 78.0%; Score 64; DB 2; Length 83;

Best Local Similarity 84.6%; Pred. No. 0.0011; 2; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14

Db 39 AWYTGSGIRPVGR 51

### RESULT 2

B70631

hypothetical protein Rv0428c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C:Accession: B70631

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

Rajandream, M.A.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Nature 393, 537-544, 1998

A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70631

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-302 <COL>

A:Cross-references: GB:Z84724; GB:AL123456; NID:G3261708; PID:CA806568.1; PID:g1817694

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv0428c

C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0428c

Query Match 53.7%; Score 44; DB 2; Length 302;  
 Best Local Similarity 70.0%; Pred. No. 8.6;  
 Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY 2 AWYASRGIRP 11  
 ||||| : :  
 Db 126 AWYASRDLP 135

RESULT 3  
 S61046  
 ARP1 protein - yeast (*Saccharomyces cerevisiae*)  
 N;Alternate names: protein D1478; protein YDL167c  
 C;Species: *Saccharomyces cerevisiae*  
 C;Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 29-Oct-1999  
 C;Accession: S61046; S31139; S67719  
 R;Pohl, T.M.  
 Submitted to the EMBL Data Library, November 1995  
 A;Reference number: S61010  
 A;Accession: S61046  
 A;Molecule type: DNA  
 A;Residues: 1-719 <POH>  
 A;Cross-references: EMBL:Z67750; NID:G1061256; PIDN:CAA91579.1; PID:G1061272  
 R;Wehner, B.P.; Rao, E.; Brendel, M.  
 Mol. Gen. Genet. 237, 351-358, 1993  
 A;Title: Molecular structure and genetic regulation of SPA, a gene responsible for resistance to the herbicide paraquat  
 A;Reference number: S31138; MUID:93247548; PMID:8483449  
 A;Accession: S31139  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-492, N', 494-719 <WEH>  
 A;Cross-references: EMBL:X68020; NID:G577609; PIDN:CAA48159.1; PID:G288530  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992  
 R;Pohl, T.M.  
 submitted to the Protein Sequence Database, July 1996  
 A;Reference number: S67708  
 A;Accession: S67719  
 A;Molecule type: DNA  
 A;Residues: 1-719 <POW>  
 A;Cross-references: EMBL:Z74215; NID:G1431265; PIDN:CAA98741.1; PID:E253076; PID:G143126  
 A;Experimental source: strain S289C  
 C;Genetics:  
 A;Gene: SGD:NRP1; ARP1  
 A;Cross-references: MIPS:YDL167c; SGD:S0002326  
 A;Map position: 4L

Query Match 53.7%; Score 44; DB 2; Length 719;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 6; Conservative 3; Mismatches 0; Gaps 0;

QY 2 AWYASRGIRPVG 13  
 : : : : :  
 Db 244 SWFTQGVVRPVG 255

RESULT 4  
 A97505  
 hypothetical protein AGR\_C\_2196 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere)  
 C;Species: *Agrobacterium tumefaciens*  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
 C;Accession: A97505  
 R;Goodner, B.; Hinkley, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*  
 A;Reference number: A97359; MUID:21608551; PMID:11743194  
 A;Accession: A97505  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-105 <KUR>  
 A;Cross-references: GB:AF007869; PIDN:AAK86994.1; PID:G15156234; GSPDB:GN00169  
 C;Genetics:

A;Gene: AGR\_C\_2196  
 A;Map position: circular chromosome

Query Match 51.2%; Score 42; DB 2; Length 105;  
 Best Local Similarity 66.7%; Pred. No. 6.8;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIR 10  
 ||||| : : : : :  
 Db 54 AWYAARGVQ 62

RESULT 5  
 AE2723  
 hypothetical protein Atull91 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)  
 C;Species: *Agrobacterium tumefaciens*  
 C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C;Accession: AE2723  
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A;Reference number: AB2577; MUID:21608550; PMID:11743193  
 A;Accession: AE2723  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-105 <KUR>  
 A;Cross-references: GB:AF008688; PIDN:AAU42203.1; PID:G17739595; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: Atull91  
 A;Map position: circular chromosome

Query Match 51.2%; Score 42; DB 2; Length 105;  
 Best Local Similarity 66.7%; Pred. No. 6.8;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 AWYASRGIR 10  
 ||||| : : : : :  
 Db 54 AWYAARGVQ 62

RESULT 6  
 AI2484  
 hypothetical protein alr7057 [imported] - *Nostoc* sp. (strain PCC 7120) plasmid pCC7120all  
 C;Species: *Nostoc* sp. PCC 7120  
 A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C;Accession: AI2484  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium Ana*  
 A;Reference number: AB1807; MUID:21595285; PMID:11759840  
 A;Accession: AI2484  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-113 <KUR>  
 A;Cross-references: GB:BA000020; PIDN:BA078141.1; PID:G17135595; GSPDB:GN00180  
 A;Experimental source: strain PCC 7120  
 C;Genetics:  
 A;Gene: alr7057  
 A;Genome: plasmid

Query Match 51.2%; Score 42; DB 2; Length 113;  
 Best Local Similarity 70.0%; Pred. No. 7.3;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AWYASRGIRP 11  
 ||||| : : : : :  
 Db 244 SWFTQGVVRPVG 255

Db 64 AWYAKAGIEP 73

RESULT 7  
B82824  
protein-export membrane protein XF0304 [imported] - Xylella fastidiosa (strain 9asc)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: B82824  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing, 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; PMID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82824  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-132 <SIM>  
A:Cross-references: GB:AE003883; GB:AE003849; NID:G9105115; PIDN:AAFE83115.1; GSPDB:GN001  
A:Experimental source: strain 9asc  
R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H. as-Neto, E.; Docena, C.; El-Borhy, H.; Pacincani, A.P.; Ferreira, A.J.S.  
submitted to Genbank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0304

Query Match 51.2%; Score 42; DB 2; Length 132;  
Best Local Similarity 61.5%; Pred. No. 8.5;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
||||||| :  
70 AWYASRGHVAQ 82

Db 70 AWYASRGHVAQ 82

RESULT 8  
G83400  
hypothetical protein PA1952 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: G83400  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bu adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path A:Reference number: A82950; PMID:20437337; PMID:10984043  
A:Accession: G83400  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-250 <STO>  
A:Cross-references: GB:AE004622; GB:AE004091; NID:G9947948; PIDN:AAG05340.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA1952

Query Match 51.2%; Score 42; DB 2; Length 250;  
Best Local Similarity 53.8%; Pred. No. 16;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CAWYASRGIRPVG 13  
||||| :|||

Db 237 CAWEQLRALRPSG 249

RESULT 9  
A95148  
licD2 protein [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
C:Accession: A95148  
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid on, J.D.; Unyama, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, nson, T.; Hickey, E.K.; White, O.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae. A:Reference number: A95000; PMID:21357209; PMID:11463916  
A:Accession: A95148  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-269 <KUR>  
A:Cross-references: GB:AE005672; PIDN:AAK75378.1; PID:G14972757; GSPDB:GN00164; TIGR:SP4 A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SPI274

Query Match 51.2%; Score 42; DB 2; Length 269;  
Best Local Similarity 54.5%; Pred. No. 17;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CAWYASRGIRP 11  
||||| :  
166 CSWYALRFVNP 176

Db 166 CSWYALRFVNP 176

RESULT 10  
G98015  
licD Protein [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001  
C:Accession: G98015  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.; A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; PMID:21429245; PMID:11544234  
A:Accession: G98015  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-269 <KUR>  
A:Cross-references: GB:AE007317; PIDN:AAK99955.1; PID:G15458781; GSPDB:GN00174 C:Genetics:  
A:Gene: licD2

Query Match 51.2%; Score 42; DB 2; Length 269;  
Best Local Similarity 54.5%; Pred. No. 17;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CAWYASRGIRP 11  
||||| :  
166 CSWYALRFVNP 176

Db 166 CSWYALRFVNP 176

RESULT 11  
E95936  
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid pSymB C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: E95936  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan Proc.Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: B95936

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-256 <KUR>

A:Cross-references: GB:AL591985; PIDN:CAC49157.1; PID:G15140642; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.

ela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMD21253

A:Genome: plasmid

Query Match 50.0%; Score 41; DB 2; Length 256;

Best Local Similarity 54.5%; Pred. No. 24;

Matches 6; Conservative 2; Mismatches 0; Indels 3; Gaps 0;

Qy 3 WYASRGIRPVG 13

Db 185 WHGTRGCRPYG 195

RESULT 12

D70817

hypothetical protein Rv1716 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C:Accession: D70817

R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, C.; Churcher, C.; Harris, D.; Gordon, S.

Rajandream, M.A.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Nature 393, 537-544, 1998

A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: D70817

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-276 <COL>

A:Cross-references: GB:AL022003; GB:AL123456; NID:G3261547; PIDN:CAA17613.1; PID:G326154

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv1716

Query Match

Best Local Similarity 50.0%; Score 41; DB 2; Length 276;

Matches 5; Conservative 5; Mismatches 0; Indels 1; Gaps 0;

Qy 3 WYASRGIRPVG 13

Db 150 WFAAKGVKAVG 160

RESULT 13

AC2368

hypothetical protein all4499 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C:Accession: AC2368

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AC2368

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-546 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA076198.1; PID:G17133635; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all4499

C:Superfamily: Synchocystis hypothetical protein slr0042

Query Match 50.0%; Score 41; DB 2; Length 546;

Best Local Similarity 66.7%; Pred. No. 49;

Matches 10; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

Qy 2 AWY--ASRGIRPVG 14

Db 457 AWYWAASLGIRDFGR 471

RESULT 14

S43878

trwC protein - Escherichia coli plasmid R388

C:Species: Escherichia coli

C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C:Accession: S43878

R:Illosa, M.; Bolland, S.; de la Cruz, F.

J. Mol. Biol. 235, 448-464, 1994

A:Title: Genetic organization of the conjugal DNA processing region of the IncW plasmid

A:Reference number: S43877; MUID:94118301; PMID:8285274

A:Accession: S43878

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-966 <LLO>

A:Cross-references: ENBL:X63150; NID:G452747; PIDN:CAA44853.1; PID:G452749

C:Genetics:

A:Genome: plasmid R388

Query Match

Best Local Similarity 50.0%; Score 41; DB 2; Length 966;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 WYASRGIRP 11

Db 242 WYAAAGLDP 250

RESULT 15

B95351

VirB6 type IV secretion protein [imported] - Sinorhizobium meliloti (strain 1021) magapi

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C:Accession: B95351

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: B95351

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-313 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK65372.1; PID:G14523833; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pega, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: virB6



C;Species: Haemophilus influenzae  
 C;Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 25-Aug-2003  
 C;Accession: I64045  
 C;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.R.; Fleischmann, R.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fink, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995  
 A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A;Reference number: A64000; MUID:95350630; PMID:7542800  
 A;Accession: I64045  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-788 <TIGR>  
 A;Cross-references: GB:U32691; GB:I42023; NID:G1573004; PIDN:AAC21739.1; PID:G1573009; TIGR:U32691.1  
 C;Genetics:  
 A;Gene: rec2  
 C;Superfamily: competence protein ComEC

Query Match 48.8%; Score 40; DB 2; Length 788;  
 Best Local Similarity 63.6%; Pred. No. 1e+02; Indels 3; Gaps 0;  
 Matches 7; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 3 WYASRGIRPVG 13  
 |||:||||  
 DB 156 WYFSKGITANG 166  
 |||:||||

RESULT 20  
 T47959  
 hypothetical protein F15G16.60 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C;Accession: T47959  
 R;De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Queb submitted to the Protein Sequence Database, January 2000  
 A;Reference number: Z24480  
 A;Accession: T47959  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-790 <DEH>  
 A;Cross-references: EMBL:AL132959  
 A;Experimental source: cultivar Columbia; BAC clone F15G16  
 C;Genetics:  
 A;Map position: 3  
 A;Introns: 39/1; 678/2; 698/3; 773/2  
 A;Note: F15G16.60

Query Match 48.8%; Score 40; DB 2; Length 790;  
 Best Local Similarity 63.6%; Pred. No. 1e+02; Indels 2; Gaps 0;  
 Matches 7; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 4 YASRGIRPVGR 14  
 |||:||||  
 DB 373 YGSRGLQPHGR 383  
 |||:||||

RESULT 21  
 F75556  
 probable acetyltransferase - Deinococcus radiodurans (strain R1)  
 C;Species: Deinococcus radiodurans  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C;Accession: F75556  
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M. Science 286, 1571-1577, 1999  
 A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A;Reference number: A75250; MUID:20036896; PMID:10567266  
 A;Accession: F75556  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-171 <WHI>

A;Cross-references: GB:AE001875; GB:AE000513; NID:G6457790; PIDN:AAF09716.1; PID:G645779 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DR0125  
 A;Map position: 1

Query Match 47.6%; Score 39; DB 2; Length 171;  
 Best Local Similarity 54.5%; Pred. No. 35; Indels 3; Gaps 0;  
 Matches 6; Conservative 2; Mismatches 3; Indels 3; Gaps 0;

QY 3 WYASRGIRPVG 13  
 |||:||||  
 DB 140 WYAAQGAREAG 150  
 |||:||||

RESULT 22  
 AE2305  
 hypothetical protein alr3996 [imported] - Nostoc sp. (strain PCC 7120)  
 C;Species: Nostoc sp. PCC 7120  
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C;Accession: AE2305  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana A;Reference number: AB1807; MUID:21595285; PMID:11759840  
 A;Accession: AE2305  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-181 <KUR>  
 A;Cross-references: GB:BA000019; PIDN:BA075695.1; PID:G17133131; GSPDB:GN00179 A;Experimental source: strain PCC 7120  
 C;Genetics:  
 A;Gene: alr3996

Query Match 47.6%; Score 39; DB 2; Length 181;  
 Best Local Similarity 58.3%; Pred. No. 37; Indels 5; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 2 AWYASRGIRPVG 13  
 |||:||||  
 DB 91 AWIAENGTPVG 102  
 |||:||||

RESULT 23  
 T35841  
 Probable membrane protein - Streptomyces coelicolor  
 C;Species: Streptomyces coelicolor  
 C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C;Accession: T35841  
 R;Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, August 1998  
 A;Reference number: Z21590  
 A;Accession: T35841  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-184 <HAR>  
 A;Cross-references: EMBL:AL031260; PIDN:CAA20292.1; GSPDB:GN000070; SCOEDB:SC9A10.05C A;Experimental source: strain A3(2)  
 C;Genetics:  
 A;Gene: SCOEDB:SC9A10.05C

Query Match 47.6%; Score 39; DB 2; Length 184;  
 Best Local Similarity 50.0%; Pred. No. 37; Indels 3; Gaps 0;  
 Matches 6; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

QY 2 AWYASRGIRPVG 13  
 |||:||||  
 DB 90 AWYSAHVRQVG 101  
 |||:||||

RESULT 24

H82925  
A:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: H82925  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
A:Reference number: A82870  
A:Accession: H82925  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-260 <GLA>  
A:Cross-references: GB:AE002116; GB:AF222894; NID:G6899118; PIDN:AAF30569.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: UU163  
A:Genetic code: SGC3

Query Match 47.6%; Score 39; DB 2; Length 260;  
Best Local Similarity 54.5%; Pred. No. 52;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 WYASRGIRPVG 13  
||||| : :  
Db 176 WYASDGVGKIG 186

RESULT 25  
A:7709  
formamidopyrimidine-DNA glycosylase CC3707 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: A87709  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
N., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: A87709  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-315 <STO>  
A:Cross-references: GB:AE005673; NID:G13425473; PIDN:AAK25669.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3707  
C:Superfamily: formamidopyrimidine-DNA glycosidase

Query Match 47.6%; Score 39; DB 2; Length 315;  
Best Local Similarity 50.0%; Pred. No. 62;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVG 13  
||||| : :  
Db 168 AWFAMGPEPLG 179

RESULT 26  
H82852  
hydroxybenzoate octaprenyltransferase XF0068 [imported] - Xylella fastidiosa (strain 9a5  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
C:Accession: H82852  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: H82852  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333 <SIM>

A:Cross-references: GB:AE003860; GB:AE003849; NID:G9104830; PIDN:AAF82881.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A.  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.  
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Prohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.N.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0068  
C:Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 47.6%; Score 39; DB 2; Length 333;  
Best Local Similarity 63.6%; Pred. No. 66;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 WYASRGIRPVG 13  
||||| : :  
Db 58 WKLARGDRPVG 68

RESULT 27  
G70947  
hypothetical protein RV3169 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: G70947  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
Connor, R.M.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: G70947  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-374 <COL>  
A:Cross-references: GB:AL021646; GB:AL123456; NID:G3242278; PIDN:CAA16634.1; PID:e124876  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: RV3169

Query Match 47.6%; Score 39; DB 2; Length 374;  
Best Local Similarity 60.0%; Pred. No. 73;  
Matches 9; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 3 WYASR---GIRPVG 13  
||||| : :  
Db 171 WLGRDRSWGIRPVG 185

RESULT 28  
B38176  
samB protein - Salmonella typhimurium  
C:Species: Salmonella typhimurium  
C:Date: 28-Aug-1992 #sequence\_revision 28-Aug-1992 #text\_change 20-Jun-2000  
C:Accession: B38176  
R:Nohmi, T.; Hakura, A.; Nakai, Y.; Watanabe, M.; Murayama, S.Y.; Sofuni, T.  
J. Bacteriol. 173, 1051-1063, 1991  
A:Title: Salmonella typhimurium has two homologous but different umuD operons: cloning  
A:Reference number: A38176; MUID:91123176; PMID:1991707  
A:Accession: B38176  
A>Status: preliminary  
A:Molecule type: DNA

A;Residues: 1-424 <NOH>  
A;Cross-references: GB|D90202; NID:g217087; PIDN:BAAL4226.1; PID:g217089  
A;Experimental source: strain LT2  
C;Genetics:  
A;Gene: samB  
C;Function:  
A;Description: restores UV mutability; involved in mutagenesis  
C;Superfamily: umuC protein  
C;Keywords: DNA repair; induced mutagenesis; SOS mutagenesis

```
Query Match      47.6%; Score 39; DB 2; Length 424;
Best Local Similarity 66.7%; Pred. No. 83;
Matches 6: Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```

Qy      3 WYASRGIRP 11
        | | | | |
Dp      390 WEAGRGIRP 398

```

RESULT 29  
AB1034  
UV protection protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (str  
A;Species: Salmonella enterica subsp. enterica serovar typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C;Accession: AB1034  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A;Reference number: AB05062; MUID:21534947; PMID:11677608  
A;Accession: AB1034  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-424 <PAR>  
A;Cross-references: GB:AL513382; PIDN:CAD06719.1; PID:gl6505372; GSPDB:GN00176  
C;Genetics:  
A;Gene: samB  
A;Superfamily: umuC protein

Query Match	47.6%;	Score 39;	DB 2;	Length 424;
Best Local Similarity	66.7%;	Pred. No. 83;		
Matches	6:	Conservative	1:	Mismatches
			2:	Indels
				Gaps
				0;

Qy 3 WYASRGIRP 11  
| : | | | |  
Db 390 WFAGRGIRP 398

RESULT 30  
JNECGT  
glycerol-3-phosphate transport protein - Escherichia coli (strain K-12)  
N;Alternate names: glycerol-3-phosphate permease  
C;Species: Escherichia coli  
C;Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 01-Mar-2002  
C;Accession: S00868; F64994  
R;Eiglmeyer, K.; Boos, W.; Cole, S.T.  
Mol. Microbiol. 1, 251-258, 1987  
A;Title: Nucleotide sequence and transcriptional startpoint of the glpT gene of Escherichia coli K-12.  
S65-phosphate transport system.  
A;Reference number: S00868; MUID:88201663; PMID:3329281  
A;Accession: S00868  
A;Molecule type: DNA  
A;Residues: 1-452 <EIG>  
A;Cross-references: EMBL:Y00536; NID:g41586; PIDN:CAA68598.1; PID:g41587  
A;Experimental source: strain K12  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, D.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97436617; PMID:9278503

A/Accession: F64994  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-452 <BLAT>  
A/Cross-references: GB:AE000314; GB:U0096; NID:gl788570; PIDN:AC75300.1; PID:gl788573;  
A/Experimental source: strain K-12, substrain MG1655

Query Match	47.6%;	Score 39;	DB 1;	Length 452;
Best Local Similarity	42.9%;	Pred. No. 88;		
Matches	6:	Conservative	1:	Mismatches
			7:	Indels
				0: Gaps
				0:

QY 1 CAWYASRGIRPVGR 14  
| | : | | |  
Db 130 CGWFOGGMGWPPCGR 143

RESULT 31

E91019  
sn-glycerol-3-phosphate permease [imported] - Escherichia coli (strain O157:H7, substrain  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: E91019  
R:Hayaishi, T.; Yaginuma, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Makino, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno  
A:Reference number: A99629; PMID:21156231; PMID:11258796  
A:Accession: E91019  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-452 <HAY>  
A:Cross-references: GB:BA000007; PID:BA36548.1; PID:gl3362595; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952

Query Match 47.6%; Score 39; DB 2; Length 452;  
Best Local Similarity 42.9%; Pred. No. 88;  
Matches 5; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 CAWYASRGIRPVGR 14  
| | : | | |  
130 CCFECOCMCWPPGGR 143

RESULT 32  
G85863  
sn-glycerol-3-phosphate permease [imported] - *Escherichia coli* (strain O157:H7, substrain  
C;Species: *Escherichia coli*  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C;Accession: G85863  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: G85863  
A;Status: preliminary



A;Molecule type: DNA  
A;Residues: 1-452 <STO>  
A;Cross-References: GB:AE005174; NID:G12516590; PIDN:AAG57371.1; GSPDB:GN00145; UWGP:Z34  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: glpT  
C;Superfamily: hexose phosphate transport protein uhpT

Query Match 47.6%; Score 39; DB 2; Length 452;  
Best Local Similarity 42.9%; Pred. No. 88;  
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 CAWYASRGIRPVGR 14  
| | | | |  
DB 130 CGWFGMGWPPGGR 143

RESULT 33  
T43558  
sulfide dehydrogenase (coenzyme Q2) (EC 1.8.5.-) precursor, mitochondrial [validated]  
C;Species: Schizosaccharomyces pombe  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
C;Accession: T43558; T40163  
R;Vande Weghe, J.G.; Ouy, D.W.  
J. Biol. Chem. 274, 13250-13257, 1999  
A;Title: A fission yeast gene for mitochondrial sulfide oxidation.  
A;Reference number: Z22577; MUID:99240711; PMID:10224084  
A;Accession: T43558  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-459 <VAN>  
A;Cross-References: EMBL:AF042283; NID:G5256827; PIDN:AAD41159.1; PID:G5256828  
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lelaure, V.; Galibert, F.  
submitted to the EMBL Data Library, October 1998  
A;Reference number: Z21908  
A;Accession: T40163  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-459 <WOO>  
A;Cross-References: EMBL:AL033385; PIDN:CAA21882.1; GSPDB:GN00067; SPDB:SPBC2G5.06C  
A;Experimental source: strain 972h-; cosmid c2G5  
C;Genetics:  
A;Gene: hmt2; SPBC2G5.06C  
A;Map position: 2  
C;Function:  
A;Description: catalyzes the reduction of quinone (coenzyme Q2) by sulfide [validated, M  
A;Note: proposed to function as a sulfide:quinone oxidoreductase  
C;Keywords: mitochondrion; oxidoreductase  
F;1-24/Domain: transit peptide (mitochondrion) #status predicted <TRNP>  
F;25-459/Product: sulfide dehydrogenase, mitochondrial #status predicted <MAT>

Query Match 47.6%; Score 39; DB 2; Length 459;  
Best Local Similarity 46.7%; Pred. No. 89;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 WYASRGIRP 11  
| | | | |  
DB 445 WYSGRGLIP 453

RESULT 34  
D84258  
alcohol dehydrogenase [imported] - Halobacterium sp. NRC-1  
C;Species: Halobacterium sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C;Accession: D84258  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S  
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A;Title: Genome sequence of Halobacterium species NRC-1.  
A;Reference number: A84160; MUID:20504483; PMID:11016950

A;Accession: D84258  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-600 <STO>  
A;Cross-References: GB:AE004437; NID:G10580574; PIDN:AAG19432.1; GSPDB:GN00138  
C;Genetics:  
A;Gene: adh3

Query Match 47.6%; Score 39; DB 2; Length 600;  
Best Local Similarity 45.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 AWYASRGIRPV 12  
| | | | |  
DB 552 AWHCTRGVFI 562

RESULT 35  
A35548  
319K protein ndvB - Rhizobium meliloti  
C;Species: Rhizobium meliloti  
C;Date: 21-Sep-1990 #sequence\_revision 21-Sep-1990 #text\_change 08-Oct-1999  
C;Accession: A35548  
R;Ielpi, L.; Dylan, T.; Ditta, G.S.; Helinski, D.R.; Stanfield, S.W.  
J. Biol. Chem. 265, 2843-2851, 1990  
A;Title: The ndvB locus of Rhizobium meliloti encodes a 319-kDa protein involved in the  
A;Reference number: A35548; MUID:90153914; PMID:2154461  
A;Accession: A35548  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-2870 <IEL>  
A;Cross-References: GB:J05219; NID:G152270; PIDN:AAA26305.1; PID:G152271  
C;Keywords: transmembrane protein

Query Match 47.6%; Score 39; DB 2; Length 2870;  
Best Local Similarity 50.0%; Pred. No. 5.2e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 WYASRGIRPV 12  
| | | | |  
DB 1126 WYRFGLEPM 1135

RESULT 36  
H72507  
hypothetical protein APE2038 - Aeropyrum pernix (strain K1)  
C;Species: Aeropyrum pernix  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
C;Accession: H72507  
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6, 83-101, 1999  
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy  
A;Reference number: A72450; MUID:99310339; PMID:10382966  
A;Accession: H72507  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-121 <KAW>  
A;Cross-References: DDBJ:AP000063; NID:G5105654; PIDN:BAAB1048.1; PID:di044834; PID:G510  
A;Experimental source: strain K1  
C;Genetics:  
A;Gene: APE2038  
C;Superfamily: Aeropyrum pernix hypothetical protein APE2038

Query Match 46.3%; Score 38; DB 2; Length 121;  
Best Local Similarity 50.0%; Pred. No. 37;  
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CAWYASRGIRPVGR 14  
| | | | |  
DB 43 CSGYARSSIRPTAR 56



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 06:37:10 ; Search time 12 Seconds  
(without alignments)  
65.088 Million cell updates/sec

Title: US-09-700-643A-7

Perfect score: 82

Sequence: 1 CAWVASRGIRPVGX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	72	87.8	87	1 PRRP_HUMAN	P81277 homo sapien
2	68	82.9	98	1 PRRP_BOVIN	P81264 bos taurus
3	64	78.0	83	1 PRRP_RAT	P81278 rattus norv
4	44	53.7	719	1 NRPI_YEAST	P32770 saccharomyc
5	43	52.4	838	1 GLGE_STRAW	Q824f0 streptomyce
6	41	50.0	402	1 EXVL_STRCO	Q94bm3 streptomyce
7	40	48.8	575	1 RECU_ERWCH	P39693 erwilia chr
8	40	48.8	580	1 ATPU_YEAST	P22136 saccharomyc
9	40	48.8	788	1 REC2_HAEIN	P44408 haemophilus
10	39	47.6	287	1 FPG_CAUCR	Q9a259 caulobacter
11	39	47.6	413	1 EXVL_COREF	Q8fcp1 corynebacte
12	39	47.6	424	1 EXVL_CORGL	Q8nm3 corynebacte
13	39	47.6	424	1 SAME_SALTY	P23832 salmonella
14	39	47.6	452	1 GLPT_ECOLI	P08194 escherichia
15	39	47.6	459	1 HMT2_SCHPO	O94284 schizosacch
16	39	47.6	2832	1 NDVB_RHIME	P20471 rhizobium m
17	38.5	47.0	570	1 TWS3_HUMAN	Q9ns93 homo sapien
18	38	46.3	220	1 TIM2_CHICK	O42146 gallus gall
19	38	46.3	294	1 CAL2_PASTE	Q94715 paramecium
20	38	46.3	306	1 CBPE_BOVIN	P00732 bos taurus
21	38	46.3	383	1 CYCR_CHRVI	O82947 chromatium
22	38	46.3	417	1 CBPE_HUMAN	P15086 homo sapien
23	38	46.3	456	1 ENO_MYCPU	Q98g50 mycoplasma
24	38	46.3	930	1 SM6C_HUMAN	Q9h3t2 homo sapien
25	37.5	45.7	415	1 BCAT_CABEL	P54688 caenorhabdi
26	37	45.1	147	1 R157_BOVIN	Q28183 bos taurus
27	37	45.1	191	1 Y064_TREPA	O83103 treponema p
28	37	45.1	257	1 DAPF_CHLTE	Q8kax9 chlorobium
29	37	45.1	359	1 ALF_HAEIN	P44429 haemophilus
30	37	45.1	397	1 ALRI_RHIL0	Q98a05 rhizobium l
31	37	45.1	430	1 SR54_HALN1	Q9hnm5 halobacteri
32	37	45.1	465	1 SR54_HALVO	Q977v2 halobacteri
33	37	45.1	482	1 NCAP_CCHFV	P27317 crinean-con

RESULT 1

PRRP\_HUMAN

ID PRRP\_HUMAN STANDARD; PRT; 87 AA.

AC P81277;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing

DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-

DE releasing peptide PrRP20].

GN PRH.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=98268781; PubMed=9607765;

RA Hiruma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,

RA Kitada C., Masuo T., Asano T., Matsumoto H., Sekiguchi M.,

RA Kurokawa T., Nishimura O., Onda H., Fujino M.;

RT "A prolactin-releasing peptide in the brain.";

RL Nature 393:272-276(1998).

[2]

TISSUE SPECIFICITY.

RX MEDLINE=99426652; PubMed=10498338;

RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,

RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,

RA Sumino Y., Fujino M.;

RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its

receptor.";

RL Regul. Pept. 83:1-10(1999).

CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the

CC expression of prolactin through its receptor GPR10. May stimulate

CC lactotrophs directly to secrete PRL.

CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

CC

CC EMBL; AB015419; BAA29027.1; -.

DR MIM; 602663; -.

DR GO; GO:0005180; F:peptide hormone; TAS.

KW Hormone; Amidation; Signal.

BY SIMILARITY.

FT SIGNAL. 1 22

FT PEPTIDE 23 53

FT PROPEP 34 53

FT PROPEP 58 87

FT MOD\_RES 53 53

FT SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;

FT AMIDATION (G-54 PROVIDE AMIDE GROUP).

FT PROLACTIN-RELEASING PEPTIDE PRRP31.

FT PROLACTIN-RELEASING PEPTIDE PRRP20.

FT

FT

FT

FT

Query Match 87.8%; Score 72; DB 1; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-06;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 AWYASRGIRPVGR 14  
 DB 40 AWYASRGIRPVGR 52

## RESULT 2

PRRP\_BOVIN STANDARD; PRT; 98 AA.  
 AC P81264;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing  
 DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-  
 GN releasing peptide PrRP20].  
 OS Bos taurus (Bovine).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Bos.  
 CX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,  
 RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,  
 RA Kurokawa T., Nishimura O., Onda H., Fujino M.,  
 RT "A prolactin-releasing peptide in the brain."  
 RL Nature 393:272-276(1998).  
 CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the  
 CC expression of prolactin through its receptor GPR10. May stimulate  
 CC lactotrophs directly to secrete PRL.  
 CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB015417; BAA29025.1;  
 CC Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
 CC SIGNAL 1 22  
 CC PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.  
 CC PROPEP 58 98 PROLACTIN-RELEASING PEPTIDE PRRP20.  
 CC MOD RES 53 53  
 CC SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 82.98; Score 68; DB 1; Length 98;  
 Best Local Similarity 92.38; Pred. No. 5e-05;  
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AWYASRGIRPVGR 14  
 DB 40 AWYASRGIRPVGR 52

## RESULT 3

PRRP\_RAT STANDARD; PRT; 83 AA.  
 AC P81278; Q8K3Y0;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing  
 DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-  
 GN releasing peptide PrRP20].  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,  
 RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,  
 RA Kurokawa T., Nishimura O., Onda H., Fujino M.,  
 RT "A prolactin-releasing peptide in the brain."  
 RL Nature 393:272-276(1998).  
 CC [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC STRAIN=Sprague-Dawley; TISSUE=Hypothalamus;  
 RA Anderson S.T., Kokay I.C., Lang T., Grattan D.R., Curlewis J.D.;  
 RT "Quantitation of prolactin-releasing peptide (PrRP) mRNA expression in  
 RT specific brain regions during the rat oestrous cycle and in  
 RT lactation."  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=99426652; PubMed=10498338;  
 RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,  
 RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,  
 RA Sumino Y., Fujino M.,  
 RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its  
 RT receptor."  
 RL Regul. Pept. 83:1-10(1999).  
 CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the  
 CC expression of prolactin through its receptor GPR10. May stimulate  
 CC lactotrophs directly to secrete PRL.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=P81278-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=P81278-2; Sequence=VSP\_004370;  
 CC -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in  
 CC medulla oblongata and hypothalamus.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB015418; BAA29026.1;  
 CC EMBL; AF521930; AAM82154.1;  
 CC PIR; JC7607; JC7607.  
 CC Hormone; Amidation; Signal; Cleavage on pair of basic residues;  
 CC Alternative splicing;  
 CC SIGNAL 1 21 BY SIMILARITY  
 CC PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRRP31.  
 CC PROPEP 57 83 PROLACTIN-RELEASING PEPTIDE PRRP20.  
 CC MOD RES 52 52  
 CC VARSPLIC 33 83  
 CC AMIDATION (G-53 PROVIDE AMIDE GROUP).  
 CC TPDIRNPAYTGRGIRPVGRRRATPRDVTGLQSLPL  
 CC DGRTKFSORG -> SECLTYGKQLTSHFTSQMPP (in  
 CC isoform 2)  
 CC /FLID=VSP\_004370.  
 CC DOC75A264BEE4F29 CRC64;  
 CC  
 CC SEQUENCE 83 AA; 9215 MW;  
 CC Score 64; DB 1; Length 83;  
 CC Best Local Similarity 84.6%; Pred. No. 0.00021;  
 CC Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



```

PT ACT_SITE 511 511 BY SIMILARITY.
FT ACT_SITE 513 513 BY SIMILARITY.
FT ACT_SITE 566 566 BY SIMILARITY.
FT ACT_SITE 633 633 BY SIMILARITY.
FT ACT_SITE 634 634 BY SIMILARITY.
SQ SEQUENCE 838 AA; 92740 MW; 3BFD6B6AD3869CE CRC64;

Query Match 52.4%; Score 43; DB 1; Length 838;
Best Local Similarity 72.7%; Pred. No. 9.8;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 AWYASRGIRPV 12
   |||:||||
Db 342 AWMAARGRPV 352
   |||:||||

RESULT 6
EX7L STRCO
ID EX7L STRCO STANDARD; PRT; 402 AA.
AC Q9PBM3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE (Exonuclease VII large subunit).
DE XSEA OR SCO5056 OR SCK7.29C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kisser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kisser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AL939122; CAC05901.1; -.
CC HAMAP; MF_00378; -.
CC InterPro; IPR003753; Exonuc VII_L.
CC InterPro; IPR004365; rRNA anti.
CC Pfam; PF02601; Exonuc VII_L; 1.
CC Pfam; PF01336; rRNA anti; 1.
CC TIGRFAMs; TIGR00237; xseA; 1.
CC Nucleotase; Nuclease; Exonuclease; Complete proteome.
CC SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;
SQ

```

```

Query Match          50.0%; Score 41; DB 1; Length 402;
Best Local Similarity 47.4%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 8; Gaps 1;

QY 3 WYASRG-----IRPVG 13
   |||||
   91 WYAPRGQLSIRAAEIKPVG 109

Db

RESULT 7
RECJ ERWCH
ID RECJ ERWCH STANDARD; PRT; 575 AA.
AC P39593; Q47007;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Single-stranded-DNA-specific exonuclease recJ (EC 3.1.1.-).
GN recJ.
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=556;
RN [1]
RP SEQUENCE FROM N.A.
RL Lovett S.T., Tzonava M., Suter V.A.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 1-100 FROM N.A.
RC MEDLINE=3937;
RX MEDLINE=94222048; PubMed=8168497;
RA Shevchik V.B., Condemine G., Robert-Baudouy J.;
RT "Characterization of Dsbc, a periplasmic protein of Erwinia
RT chrysanthemi and Escherichia coli with disulfide isomerase
RT activity.";
RL EMBO J. 13:2007-2012(1994).
CC -I- FUNCTION: SINGLE-STRANDED-DNA-SPECIFIC EXONUCLEASE. REQUIRED FOR
CC MANY TYPES OF RECOMBINATIONAL EVENTS. ALTHOUGH THE STRINGENCY OF
CC THE REQUIREMENT FOR RECJ APPEARS TO VARY WITH THE TYPE OF
CC RECOMBINATIONAL EVENT MONITORED AND THE OTHER RECOMBINATION GENE
CC PRODUCTS WHICH ARE AVAILABLE [BY SIMILARITY].
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U57963; AAB02261.1; -
DR EMBL; X76687; CAA54109.1; -
DR InterPro; IPR003156; DHH1.
DR InterPro; IPR001667; Ppsterase.
DR InterPro; IPR004610; RecJ.
DR Pfam; PF01368; DHH; 1.
DR Pfam; PF02272; DHH1; 1.
DR TIGRfams; TIGR00644; recJ; 1.
KW Hydroxylase; Nuclease; Exonuclease.
FT CONFLICT 1 6 MVVTVQ -> MLLPN (IN REF. 2).
FT CONFLICT 83 83 R -> G (IN REF. 2).
FT CONFLICT 100 100 R -> H (IN REF. 2).
SQ SEQUENCE 575 AA; 62924 MW; 4D06A037DFD9DBE4 CRC64;

Query Match          48.8%; Score 40; DB 1; Length 575;
Best Local Similarity 55.6%; Pred. No. 23;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAWYASRGI 9
   | | | | |
Db 208 CGWFAERGL 216

```

```

Query Match          50.0%; Score 41; DB 1; Length 402;
Best Local Similarity 47.4%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 8; Gaps 1;

QY 3 WYASRG-----IRPVG 13
   |||||
   91 WYAPRGQLSIRAAEIKPVG 109

Db

RESULT 7
RECJ ERWCH
ID RECJ ERWCH STANDARD; PRT; 575 AA.
AC P39593; Q47007;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Single-stranded-DNA-specific exonuclease recJ (EC 3.1.1.-).
GN recJ.
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=556;
RN [1]
RP SEQUENCE FROM N.A.
RL Lovett S.T., Tzonava M., Suter V.A.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 1-100 FROM N.A.
RC MEDLINE=3937;
RX MEDLINE=94222048; PubMed=8168497;
RA Shevchik V.B., Condemine G., Robert-Baudouy J.;
RT "Characterization of Dsbc, a periplasmic protein of Erwinia
RT chrysanthemi and Escherichia coli with disulfide isomerase
RT activity.";
RL EMBO J. 13:2007-2012(1994).
CC CC -1- FUNCTION: SINGLE-STRANDED-DNA-SPECIFIC EXONUCLEASE. REQUIRED FOR
CC MANY TYPES OF RECOMBINATIONAL EVENTS. ALTHOUGH THE STRINGENCY OF
CC THE REQUIREMENT FOR RECJ APPEARS TO VARY WITH THE TYPE OF
CC RECOMBINATIONAL EVENT MONITORED AND THE OTHER RECOMBINATION GENE
CC PRODUCTS WHICH ARE AVAILABLE [BY SIMILARITY].
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U57963; AAB02261.1; -
DR EMBL; X76687; CAA54109.1; -
DR InterPro; IPR003156; DHHA1.
DR InterPro; IPR001667; Ppsterase.
DR InterPro; IPR004610; RecJ.
DR Pfam; PF01368; DHH; 1.
DR Pfam; PF02272; DHHA1; 1.
DR TIGRFAMs; TIGR00644; recJ; 1.
KW Hydroxylase; Nuclease; Exonuclease.
FT CONFLICT 1 6 MVVVTQ -> MLLPN (IN REF. 2).
FT CONFLICT 83 83 R -> G (IN REF. 2).
FT CONFLICT 100 100 R -> H (IN REF. 2).
SQ SEQUENCE 575 AA; 62924 MW; 4D06A037DFD9DBE4 CRC64;

Query Match          48.8%; Score 40; DB 1; Length 575;
Best Local Similarity 55.6%; Pred. No. 23;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAWYASRGI 9
   | | | | |
Db 208 CGWFAERGL 216

```

```

RESULT 8
ATPU YEAST
ID ATPU_YEAST STANDARD; PRT; 580 AA.
AC P22136; P23115;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ATP13 protein, mitochondrial precursor.
GN ATP13 OR AEP2 OR YMR282C OR YMR021.08C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91122292; PubMed=1825065;
RA Ackerman S.H., Gatti D.L., Gellefors P., Douglas M.G., Tzagoloff A.;
RT "ATP13, a nuclear gene of Saccharomyces cerevisiae essential for the
RL expression of subunit 9 of the mitochondrial ATPase."
RN FEBS Lett. 278:234-238(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92035073; PubMed=1718609;
RA Payne M.J., Finnegan P.M., Keramidaris E., Lukins H.B.;
RT "Characterization of a yeast nuclear gene, AEP2, required for
RL accumulation of mitochondrial mRNA encoding subunit 9 of the ATP
RN synthase."
RN Curr. Genet. 20:53-61(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RN XIII."
RL Nature 387:90-93(1997).
RN [4]
RP SEQUENCE OF 1-108 FROM N.A.
RX STRAIN=S288c / YPH1;
RX MEDLINE=95042750; PubMed=7954819;
RA Aastrom S.U., Byström A.S.;
RT "Ril1, a tRNA backbone-modifying enzyme that mediates initiator and
RN elongator tRNA discrimination."
RL Cell 79:1535-1546(1994).
CC -!- FUNCTION: Essential for the expression of subunit 9 of the
CC mitochondrial ATPase.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- MISCELLANEOUS: Mutations in ATP13 result in respiratory
CC deficiency.
CC -!- CAUTION: Ref.1 sequence differs from that shown from position 363
CC onward and is shorter (372 AA) due to a frameshift.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56215; CAA39672.2; ALT_FRAME.
CC EMBL; M59860; AAA34412.1; -.
CC EMBL; Z49704; CAA89780.1; -.
CC EMBL; X80795; CAA56771.1; -.
CC PIR; S54589; S54589.
CC SGD; S0004895; AEP2.
CC Mitochondrion; Transit peptide.
CC TRANSIT 1 25 MITOCHONDRION (OR 28) (POTENTIAL).
CC CHAIN 26 580 ATP13 PROTEIN.

```

---

```

FT CONFLICT 214 214 G -> N (IN REF. 1).
FT CONFLICT 481 481 I -> V (IN REF. 1).
SQ SEQUENCE 580 AA; 67523 MW; BC0AB10EABCD44AA CRC64;

Query Match 48.8%; Score 40; DB 1; Length 580;
Best Local Similarity 66.7%; Pred.No. 23;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 3 WYASRGIRP 11
Db 439 WYASKRLRP 447
||||: |||

RESULT 9
REC2 HAEIN STANDARD; PRT; 788 AA.
ID REC2_HAEIN
AC P44408;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Recombination protein 2.
GN REC2 OR REC-2 OR HI0061.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BC200;
RX MEDLINE=94341577; PubMed=8063112;
RA Clifton S.W., McCarthy D., Roe B.A.;
RT "Sequence of the rec-2 locus of Haemophilus influenzae: homologies to
RN comE-CRP3 of Bacillus subtilis and msbA of Escherichia coli."
RL Gene 146:95-100(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Kerslavage A.R., Bult C.J., Tomb J.-P., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RN Rd."
RL Science 269:496-512(1995).
CC -!- FUNCTION: Might contribute to transformation as a member of a
CC membrane-bound pore complex at the base of the transformosome. It
CC could directly interact with transforming DNA during translocation
CC indirectly by participating in the assembly of the pore.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -!- SIMILARITY: TO B.SUBTILIS COMEC, N.GONORRHOEA COMA, AND E.COLI
CC YCAI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L20805; AAC13733.1; -.
CC EMBL; U32691; AAC21739.1; -.
CC PIR; I64045; I64045.
CC TIGR; HI0061; -.
CC InterPro; IPR001279; Blactmase-like.
CC InterPro; IPR004477; COMEC_N-term.

```





DR InterPro; IPR008994; Nucleic\_acid\_OB.  
DR Pfam; PF02601; Exonuc\_VII\_L; 1.  
DR TIGRFAMs; TIGR00237; xsea; 1.  
KW Hydroxylase; Nuclease; Exonuclease; Complete proteome.  
SQ SEQUENCE 413 AA; 45171 MW; D3BABD9687C6EE5D CRC64;  
  
Query Match 47.6%; Score 39; DB 1; Length 413;  
Best Local Similarity 50.0%; Pred. No. 25;  
Matches 10; Conservative 1; Mismatches 1; Indels 8; Gaps 1;  
  
QY 2 AWYASRG-----IRPVG 13  
|:||||  
DB 91 AFVAGRGTSFWTDIRPVG 110

RESULT 12  
EXYL\_CORGL STANDARD; PRT; 417 AA.

ID EXYL\_CORGL AC Q8NRM3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)  
DE [Exonuclease VII large subunit]  
GN XSEA OR CGI1025.  
OS Corynebacterium glutamicum (Brevibacterium flavum).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=1719;  
[1]  
RP SEQUENCE FROM N.A.  
RN STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
RA Nakagawa S.;  
RC Submitted (May-2002) to the EMBL/GenBank/DDBJ databases.  
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
RL FUNCTION: Bidirectionally degrades single-stranded DNA into large acid-insoluble oligonucleotides, which are then degraded further into small acid-soluble oligonucleotides (By similarity).  
CC -! CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5' to 3' or 3' to 5'-direction to yield nucleoside 5'-phosphates.  
CC -! SUBUNIT: Heterooligomer composed of large and small subunits (By similarity).  
CC -! SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -! SIMILARITY: Belongs to the xsea family.  
-----  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).  
-----  
EMBL; APC05277; BAB98418.1; --  
HAMAP; MF\_00378; -, 1.  
InterPro; IPR003753; Exonuc\_VII\_L.  
DR InterPro; IPR008994; Nucleic\_acid\_OB.  
DR Pfam; PF02601; Exonuc\_VII\_L; 1.  
DR TIGRFAMs; TIGR00237; xsea; 1.  
KW Hydroxylase; Nuclease; Exonuclease; Complete proteome.  
SQ SEQUENCE 417 AA; 45582 MW; B3ACD9286C173C34 CRC64;  
  
Query Match 47.6%; Score 39; DB 1; Length 417;  
Best Local Similarity 50.0%; Pred. No. 25;  
Matches 10; Conservative 1; Mismatches 1; Indels 8; Gaps 1;  
  
QY 2 AWYASRG-----IRPVG 13  
|:||||  
DB 95 AFVAGRGTSFWTDIRPVG 114

RESULT 13  
SAMB\_SALTY STANDARD; PRT; 424 AA.

ID SAMB\_SALTY AC P23832;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE SamB protein.  
GN SAMB.  
OS Salmonella typhimurium.  
OG Plasmid 60-mDa cryptic.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=602;  
[1]  
RP SEQUENCE FROM N.A.  
RN STRAIN=LT2;  
RC MEDLINE=91123176; PubMed=1991707;  
RA Nohmi T., Hakura A., Nakai Y., Watanabe M., Murayama S.Y.,  
RA Sofuni T.;  
RT "Salmonella typhimurium has two homologous but different umuDc operons: cloning of a new umuDc-like operon (samAb) present in a 60-megadalton cryptic plasmid of S. typhimurium.";  
RL J. Bacteriol. 173:1051-1063(1991).  
CC -! FUNCTION: Involved in UV protection and mutation.  
CC -! SIMILARITY: Belongs to the DNA polymerase type-Y family.  
CC -! SIMILARITY: Contains 1 umuC domain.  
-----  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).  
-----  
EMBL; D90202; BAAL4226.1; --  
FIR; B38176; B38176.  
HAMAP; MF\_01113; atypical; 1.  
InterPro; IPR001126; UMUC\_like.  
DR Pfam; PF00817; IMS; 1.  
DR PROSITE; PS0173; UMUC; 1.  
KW Plasmid; SOS mutagenesis; DNA repair.  
FT DOMAIN 2 189 UMUC  
SQ SEQUENCE 424 AA; 47727 MW; FF8C47476CC58A2B CRC64;  
  
Query Match 47.6%; Score 39; DB 1; Length 424;  
Best Local Similarity 66.7%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 3 WYASRGIRP 11  
|:||||  
DB 390 WFAGRGIAP 398

RESULT 14  
GLPT\_ECOLI STANDARD; PRT; 452 AA.

ID GLPT\_ECOLI AC P08194;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Glycerol-3-phosphate transporter (G-3-P transporter) (G-3-P permease).  
GN GLPT OR B2240.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
[1]  
RP SEQUENCE FROM N.A.  
RN STRAIN=KL2;  
RC MEDLINE=88201663; PubMed=3329281;  
RX Eiglmeyer K., Boos W., Cole S.;  
RT "Nucleotide sequence and transcriptional startpoint of the glpt gene"

DR InterPro; IPR008994; Nucleic\_acid\_OB.  
DR Pfam; PF02601; Exonuc\_VII\_L; 1.  
DR TIGRFAMs; TIGR00237; xsea; 1.  
KW Hydroxylase; Nuclease; Exonuclease; Complete proteome.  
SQ SEQUENCE 413 AA; 45171 MW; D3BABD9687C6EE5D CRC64;  
  
Query Match 47.6%; Score 39; DB 1; Length 413;  
Best Local Similarity 50.0%; Pred. No. 25;  
Matches 10; Conservative 1; Mismatches 1; Indels 8; Gaps 1;  
  
QY 2 AWYASRG-----IRPVG 13  
|:||||  
DB 91 AFVAGRGTSFWTDIRPVG 110

RESULT 12  
EXYL\_CORGL STANDARD; PRT; 417 AA.

ID AC Q8NRM3;  
DT DT 28-FEB-2003 (Rel. 41, Created)  
DT DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)  
DE DE (Exonuclease VII large subunit).  
GN XSEA OR CGI1025.  
OS Corynebacterium glutamicum (Brevibacterium flavum).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=1719;  
[1]  
RP SEQUENCE FROM N.A.  
RN STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
RA Nakagawa S.;  
RC Submitted (May-2002) to the EMBL/GenBank/DDBJ databases.  
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
RL FUNCTION: Bidirectionally degrades single-stranded DNA into large  
acid-insoluble oligonucleotides, which are then degraded further  
into small acid-soluble oligonucleotides (By similarity).  
CC -! CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5' to 3'-  
or 3' to 5'-direction to yield nucleoside 5'-phosphates.  
CC -! SUBUNIT: Heterooligomer composed of large and small subunits (By  
similarity).  
CC -! SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -! SIMILARITY: Belongs to the xsea family.  
-----  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announcement/  
or send an email to license@isb-sib.ch).  
-----  
EMBL; APC05277; BAB98418.1; --  
HAMAP; MF\_00378; -, 1.  
InterPro; IPR003753; Exonuc\_VII\_L.  
DR InterPro; IPR008994; Nucleic\_acid\_OB.  
DR Pfam; PF02601; Exonuc\_VII\_L; 1.  
DR TIGRFAMs; TIGR00237; xsea; 1.  
KW Hydroxylase; Nuclease; Exonuclease; Complete proteome.  
SQ SEQUENCE 417 AA; 45582 MW; B3ACD9286C173C34 CRC64;  
  
Query Match 47.6%; Score 39; DB 1; Length 417;  
Best Local Similarity 50.0%; Pred. No. 25;  
Matches 10; Conservative 1; Mismatches 1; Indels 8; Gaps 1;  
  
QY 2 AWYASRG-----IRPVG 13  
|:||||  
DB 95 AFVAGRGTSFWTDIRPVG 114

RESULT 13  
SAMB\_SALTY STANDARD; PRT; 424 AA.

ID AC P23832;  
DT DT 01-NOV-1991 (Rel. 20, Created)  
DT DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE DE SamB protein.  
GN SAMB.  
OS Salmonella typhimurium.  
OG Plasmid 60-mDa cryptic.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=602;  
[1]  
RP SEQUENCE FROM N.A.  
RN STRAIN=LT2;  
RC MEDLINE=91123176; PubMed=1991707;  
RA Nohmi T., Hakura A., Nakai Y., Watanabe M., Murayama S.Y.,  
RA Sofuni T.;  
RT "Salmonella typhimurium has two homologous but different umuDc  
operons: cloning of a new umuDc-like operon (samAb) present in a  
60-megadalton cryptic plasmid of S. typhimurium.";  
RL J. Bacteriol. 173:1051-1063(1991).  
CC -! FUNCTION: Involved in UV protection and mutation.  
CC -! SIMILARITY: Belongs to the DNA polymerase type-Y family.  
CC -! SIMILARITY: Contains 1 umuC domain.  
-----  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announcement/  
or send an email to license@isb-sib.ch).  
-----  
EMBL; D90202; BAAL4226.1; --  
FIR; B38176; B38176.  
HAMAP; MF\_01113; atypical; 1.  
InterPro; IPR001126; UMUC\_like.  
DR Pfam; PF00817; IMS; 1.  
DR PROSITE; PS0173; UMUC; 1.  
KW Plasmid; SOS mutagenesis; DNA repair.  
FT DOMAIN 2 189  
SQ SEQUENCE 424 AA; 47727 MW; FF8C47476CC58A2B CRC64;  
  
Query Match 47.6%; Score 39; DB 1; Length 424;  
Best Local Similarity 66.7%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 3 WYASRGIRP 11  
|:||||  
DB 390 WFAGRGIAP 398

RESULT 14  
GLPT\_ECOLI STANDARD; PRT; 452 AA.

ID AC P08194;  
DT DT 01-AUG-1988 (Rel. 08, Created)  
DT DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE DE Glycerol-3-phosphate transporter (G-3-P transporter) (G-3-P  
permease).  
GN GLPT OR B2240.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
[1]  
RP SEQUENCE FROM N.A.  
RN STRAIN=KL12;  
RC MEDLINE=88201663; PubMed=3329281;  
RX Eiglmeyer K., Boos W., Cole S.;  
RT "Nucleotide sequence and transcriptional startpoint of the glpt gene

of *Escherichia coli*: extensive sequence homology of the glycerol-3-phosphate transport protein with components of the hexose-6-phosphate transport system.";

Mol. Microbiol. 1:251-258 (1987).

[2]

SEQUENCE FROM N.A.

STRAIN-K12 / MG1655;

MEDLINE=97426617; PubMed=9278503;

Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

Mau B., Shao Y.;

"The complete genome sequence of *Escherichia coli* K-12.";

Science 277:1453-1474 (1997).

[3]

SEQUENCE FROM N.A.

STRAIN-K12;

MEDLINE=97349980; PubMed=9205837;

Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Iseno K.,

Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,

Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,

Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubramanian S.,

Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,

Yamagata S., Horiuchi T.;

"Construction of a contiguous 874-kb sequence of the *Escherichia coli*

- K12 genome corresponding to 50.0-68.8 min on the linkage map and

analysis of its sequence features.";

DNA Res. 4:91-113 (1997).

[4]

TOPOLOGY.

MEDLINE=89039254; PubMed=3141744;

Goett P., Boos W.;

"The transmembrane topology of the sn-glycerol-3-phosphate permease of

*Escherichia coli* analysed by phoA and lacZ protein fusions.";

Mol. Microbiol. 2:655-663 (1988).

[5]

- FUNCTION: Responsible for glycerol-3-phosphate uptake.

- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

- SIMILARITY: Belongs to the major facilitator (MFS) superfamily.

SUC37A family.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; Y00536; CAA68598.1; -

EMBL; AE000314; AAC75300.1; -

EMBL; D90855; BAA16059.1; -

PIR; S00868; JNECGT.

Ecogene; EG10401; Glpt.

InterPro; IPR005267; Glpt transporter.

InterPro; IPR000849; Glpt transporter.

InterPro; IPR007114; MFS.

TIGRFAMs; TIGR00881; 2A0104; 1.

PROSITE; TIGR00712; glpt; 1.

PROSITE; PS00942; GLPT; 1.

PROSITE; PS00850; MFS; 1.

Glycerol metabolism; Transmembrane; Inner membrane; Transport; Complete proteome.

DOMAIN 1 27

TRANSMEM 28 44

DOMAIN 45 64

TRANSMEM 65 87

DOMAIN 88 95

TRANSMEM 96 113

DOMAIN 114 119

TRANSMEM 120 136

DOMAIN 137 169

TRANSMEM 170 184

DOMAIN 185 190

PERIPLASMIC (PROBABLE).

1 (PROBABLE).

2 (PROBABLE).

3 (PROBABLE).

4 (PROBABLE).

5 (PROBABLE).

PERIPLASMIC (PROBABLE).

PERIPLASMIC (PROBABLE).

PERIPLASMIC (PROBABLE).

PERIPLASMIC (PROBABLE).

PERIPLASMIC (PROBABLE).

PERIPLASMIC (PROBABLE).

PERIPLASMIC (PROBABLE).

FT TRANSMEM 191 209  
 FT DOMAIN 210 255  
 FT TRANSMEM 256 270  
 FT DOMAIN 271 292  
 FT TRANSMEM 293 310  
 FT DOMAIN 311 323  
 FT TRANSMEM 324 348  
 FT DOMAIN 349 355  
 FT TRANSMEM 356 378  
 FT DOMAIN 379 384  
 FT TRANSMEM 385 410  
 FT DOMAIN 411 418  
 FT TRANSMEM 419 438  
 FT DOMAIN 439 452  
 SQ SEQUENCE 452 AA; 50310 MW; 0DE70D08D40AD445 CRC64;

Query Match 47.6%; Score 39; DB 1; Length 452;

Best Local Similarity 42.9%; Pred. No. 27;

Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 CAWVASRGIRPVGR 14

DB 130 CQWFGQMGWPPGGR 143

RESULT 15

HMT2\_SCHPO

ID HMT2 SCHPO STANDARD; PRT; 459 AA.

AC 094284; O13293;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Sulfide:quinone oxidoreductase, mitochondrial precursor (EC 1.1.1.1)

DE (Heavy metal tolerance protein 2) (Cadmium resistance protein 1).

GN HMT2 OR CAD1 OR SPBC2G5.06C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OC NCBI\_TaxID=4896;

[1]

SEQUENCE FROM N.A., FUNCTION, COFACTOR, SUBCELLULAR LOCATION, AND

VARIANT ANALYSIS.

RA MEDLINE=99240711; PubMed=10224084;

RA Vande Weghe J.G.; Ow D.W.;

RT "A fission yeast gene for mitochondrial sulfide oxidation.";

RL J. Biol. Chem. 274:13250-13257 (1999).

[2]

SEQUENCE FROM N.A.

RA Mutoh N., Kawabata M., Nakagawa C., Yamada K.;

RT "Molecular cloning of the gene involved in cadmium sensitivity of

fission yeast *Schizosaccharomyces pombe*.";

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

[3]

SEQUENCE FROM N.A.

RA MEDLINE=21848401; PubMed=11859360;

RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

Brooks K., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

Hollroyd S., Goble A., Hamlin N., Harris D., Hidayat J., Hodgson G.,

James K., Jones L., Howarth S., Huckle E.J., Hunt S., Jagels K.,

Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,

Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,

Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,

Borsym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
CC  
CC -!- FUNCTION: Catalyzes the reduction, with the use of NADP, of  
CC hydrogen sulfide into sulfite and water. This process confers  
CC resistance to heavy metals such as cadmium and cisplatin.  
CC -!- COFACTOR: FAD; noncovalently bound.  
CC -!- SUBCELLULAR LOCATION: Mitochondrial.  
CC -!- CAUTION: Ref.2 sequence differs from that shown extensively due to  
CC framehifts and other sequencing errors.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; J05219; AAA26305.1; ALT INIT.  
CC EMBL; AL591793; CAC47865.1; --  
CC PIR; A35548; A35548.  
CC InterPro; IPR008928; Glyco trans 6hp.  
CC Transmembrane; inner membrane; Complete proteome.  
CC TRANSMEM 411 431 POTENTIAL.  
CC TRANSMEM 444 464 POTENTIAL.  
CC TRANSMEM 810 830 POTENTIAL.  
CC TRANSMEM 831 851 POTENTIAL.  
CC TRANSMEM 959 979 POTENTIAL.  
CC CONFLICT 47 47 T -> A (IN REF. 1).  
CC CONFLICT 1187 1187 L -> F (IN REF. 1).  
CC CONFLICT 2602 2602 D -> G (IN REF. 1).  
CC FT CONFLICT 2602 2602  
CC SQ SEQUENCE 2832 AA; 315746 MW; 171EA89F03A936F2 CRC64;  
  
Query Match 47.6%; Score 39; DB 1; Length 459;  
Best Local Similarity 56.7%; Pred. No. 27;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 3 WYASRGIRP 11  
Db 445 WYGRGLIP 453  
  
RESULT 16  
ID NDVB RHIME STANDARD; PRT; 2832 AA.  
AC P20471;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Protein ndvb.  
GN NDVB OR R03286 OR SMC04382.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90153914; PubMed=2154461;  
RA Telpi L., Dylan T., Ditta G.S., Helinski D.R., Stanfield S.W.;  
RT "The ndvb locus of Rhizobium meliloti encodes a 319-kDa protein  
RT involved in the production of beta-(1-->2)-glucan.";  
RL J. Biol. Chem. 265:2843-2851(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21396507; PubMed=11481430;

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,  
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,  
RT "Analysis of the chromosome sequence of the legume symbiont  
RT Sinorhizobium meliloti strain 1021.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
CC  
CC -!- FUNCTION: INVOLVED IN THE PRODUCTION OF BETA-(1,2)-GLUCAN.  
CC IT IS INVOLVED NOT ONLY IN INVASION BUT ALSO IN BACTEROID  
CC DEVELOPMENT.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
CC -!- SIMILARITY: TO A.TUMEFACIENS CHVB.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; J05219; AAA26305.1; ALT INIT.  
CC EMBL; AL591793; CAC47865.1; --  
CC PIR; A35548; A35548.  
CC InterPro; IPR008928; Glyco trans 6hp.  
CC Transmembrane; inner membrane; Complete proteome.  
CC TRANSMEM 411 431 POTENTIAL.  
CC TRANSMEM 444 464 POTENTIAL.  
CC TRANSMEM 810 830 POTENTIAL.  
CC TRANSMEM 831 851 POTENTIAL.  
CC TRANSMEM 959 979 POTENTIAL.  
CC CONFLICT 47 47 T -> A (IN REF. 1).  
CC CONFLICT 1187 1187 L -> F (IN REF. 1).  
CC CONFLICT 2602 2602 D -> G (IN REF. 1).  
CC FT CONFLICT 2602 2602  
CC SQ SEQUENCE 2832 AA; 315746 MW; 171EA89F03A936F2 CRC64;  
  
Query Match 47.6%; Score 39; DB 1; Length 2832;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 3 WYASRGIRP 12  
Db 1088 WYTRGLEPM 1097  
  
RESULT 17  
ID T7S3 HUMAN STANDARD; PRT; 570 AA.  
AC Q9NS93; Q9NUS4;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Transmembrane 7 superfamily protein member 3 precursor (Seven span  
DE transmembrane protein).  
GN TM7SF3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20291015; PubMed=10828615;  
RA Akashi H., Han H.-J., Iizaka M., Nakajima Y., Furukawa Y., Sugano S.,  
RA Imai K., Nakamura Y.;  
RT "Isolation and characterization of a novel gene encoding a putative  
RT seven-span transmembrane protein, TM7SF3.";  
RL Cytogenet. Cell Genet. 88:305-309(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,

Wegatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
Matatabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
Yanamoto J., Wakamatsu A., Nakamura Y., Nagahazi K., Masuho Y.,  
Ninomiya K., Iwayanagi T.,  
"NEDO human cDNA sequencing project";  
Submitted (FEB-2000) to the EMBL/GenBank/DBAJ databases.

[3]

SEQUENCE FROM N.A.  
TISSUE=Placenta;  
MEDLINE=22388257; Pubmed=12477932;  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Hang J., Haieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shchevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
"Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

!- SUBCELLULAR LOCATION: Integral membrane protein; plasma membrane.

!- TISSUE SPECIFICITY: Ubiquitous. The highest expression is in  
kidney.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AB032470; BRA92856.1; -;  
EMBL; AK002031; BRA92046.1; -;  
EMBL; BC005176; AAH05176.1; -;  
Genew: HGNC:23049; TW7SF3.  
MIM; 605181; -;  
Signal; Transmembrane.  
SIGNAL 1 21  
CHAIN 22 570

POTENTIAL.  
TRANSMEMBRANE 7 SUPERFAMILY PROTEIN  
MEMBER 3.

TRANSMEM 296 313 POTENTIAL.  
TRANSMEM 320 342 POTENTIAL.  
TRANSMEM 347 369 POTENTIAL.  
TRANSMEM 371 393 POTENTIAL.  
TRANSMEM 408 430 POTENTIAL.  
TRANSMEM 437 459 POTENTIAL.  
TRANSMEM 479 501 POTENTIAL.  
TRANSMEM 297 332 PHE-RICH.  
DOMAIN 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 75 75 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
CARBOHYD 264 264 N-LINKED (GLCNAC. .) (POTENTIAL).  
CONFLICT 564 564 E -> A (IN REF. 2).  
SEQUENCE 570 AA; 64166 MW; CC296D7C22AD894C CRC64;

Query Match 47.0%; Score 38.5; DB 1; Length 570;  
Best Local Similarity 53.8%; Pred. No. 41;  
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

1 CAWY-ASRGIRPV 12

```

Db      108 CWTWLG7SGIQPV 120
||||| : |||||
RESULT 18
TIM2_CHICK
TIM2_ID TIM2_CHICK STANDARD; PRT; 220 AA.
O42146;
DT DT 15-JUL-1998 (Rel. 36, Created)
DT DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE DE Metalloproteinase inhibitor 2 precursor (TIMP-2) (Tissue inhibitor of
DE DE metalloproteinases-2).
TIMP2.
GN GN Gallus gallus (Chicken).
OS OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC OC Gallus.
OX OX NCBI_TaxID=9031;
[1]
SEQUENCE FROM N.A.
MEDLINE=98122528; PubMed=9462696;
Aimes R.T., Li L.H., Weaver B., Hawkes S., Hahn-Dantona E.A.,
Quigley J.P.;
RA "Cloning, expression, and characterization of chicken tissue
RT inhibitor of metalloproteinase-2 (TIMP-2) in normal and transformed
RT chicken embryo fibroblasts."; (1998).
RT J. Cell. Physiol. 174:342-352(1998).
RT CC -!- FUNCTION: Complexes with metalloproteinases (such as collagenases)
RT CC and irreversibly inactivates them (By similarity).
RT CC -!- SUBCELLULAR LOCATION: Secreted.
RT CC -!- PTM: The activity of TIMP2 is dependent on the presence of
RT CC disulfide bonds.
RT CC -!- SIMILARITY: Belongs to the TIMP family.
RT CC -!- SIMILARITY: Contains 1 NTR domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF004664; AAB69168.1; -.
DR HSPD; P16035; 2TMP.
DR InterPro; IPR001820; TIMP.
DR InterPro; IPR008993; TIMP_like.
DR Pfam; PF00365; TIMP; 1.
DR SMART; SM00206; NTR; 1.
DR PROSITE; PS00189; NTR; 1.
DR PROSITE; PS00288; TIMP; 1.
DR Metalloprotease inhibitor; Signal.
KW SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 220 METALLOPROTEINASE INHIBITOR 2.
FT DOMAIN 27 152 NTR.
FT DISULFID 27 98 BY SIMILARITY.
FT DISULFID 29 127 BY SIMILARITY.
FT DISULFID 39 152 BY SIMILARITY.
FT DISULFID 154 201 BY SIMILARITY.
FT DISULFID 159 164 BY SIMILARITY.
FT DISULFID 172 193 BY SIMILARITY.
SQ SEQUENCE 220 AA; 24313 MW; 61BDAC760E752E53 CRC64;
Query Match 46.3%; Score 38; DB 1; Length 220;
Best Local Similarity 63.6%; Pred No; 20;
Matches 7; Conservative 1; Mismatches 1; Indels 2; Gaps 1;
QY 1 CAWYASRGIRP 11
||||| : |||||
Db 201 CAWY--RGWAP 209

```

```

RESULT 19
CAL2_PARTE
ID -CAL2_PARTE STANDARD; PRT; 294 AA.
AC Q9475;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative cathepsin L2 (EC 3.4.22.15) (Fragment).
OS Paramacium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
OC Paramacium.
OX NCBI_TaxID=5888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Stock 51;
RX MEDLINE=96248439; PubMed=8665938;
RA Voelkel H., Kurz U., Linder J., Klumpp S., Gnaul V., Jung G.,
RA Schultz J.E.;
RT "Cathepsin L is an intracellular and extracellular protease in
RT Paramacium tetraurelia: Purification, cloning, sequencing and specific
RT inhibition by its expressed propeptide.";
RL Eur. J. Biochem. 238:198-206(1996).
CC -!- FUNCTION: May be involved in extracellular digestion.
CC -!- CATALYTIC ACTIVITY: Specificity close to that of papain. As
CC compared to cathepsin B, cathepsin L exhibits higher activity
CC towards protein substrates, but has little activity on Z-Arg-Arg-
CC NMEc, and no peptidyl-dipeptidase activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to peptidase family C1.
CC -!- CAUTION: This protein may be non-functional as it lacks the
CC cysteine active site residue which is replaced by Gly-118.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; X91756; CAA62871.1; -
DR PIR; S68784; S68784.
DR HSP; O60911; 1FH0.
DR InterPro; IPR000668; Peptidase C1.
DR InterPro; IPR000169; SHprot acsite.
DR Pfam; PF00112; Peptidase_C1_1.
DR Prodom; PD000158; Peptidase_C1; 1.
DR SMART; SM00645; Pept_C1; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE_NEG.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; FALSE_NEG.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE_NEG.
KW Hydrolase; Thiol protease.
FT NON_TER 1
FT PROPEP <1 ? ACTIVATION PEPTIDE.
FT CHAIN ? 294 PUTATIVE CATHEPSIN L2.
FT SITE 118 118 POSITION OF ACTIVE SITE CYS.
FT ACT_SITE 247 247 BY SIMILARITY.
FT ACT_SITE 264 264 BY SIMILARITY.
FT ACT_SITE 115 156 BY SIMILARITY.
FT DISULFID 240 284 BY SIMILARITY.
FT DISULFID 294 AA; 33627 MW; 6FD686AC176BEC51 CRC64;
SQ SEQUENCE 294 AA; 33627 MW; 6FD686AC176BEC51 CRC64;

Query Match 46.3%; Score 38; DB 1; Length 294;
Best Local Similarity 54.5%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ANYASRGIRPV 12
|||||
DB 129 ANYAIRGQKI 139

RESULT 20
CBPB_BOVIN
ID CBPB_BOVIN STANDARD; PRT; 306 AA.
AC P00732;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carboxypeptidase B (EC 3.4.17.2).
GN CPB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=75217824; PubMed=1057162;
RA Titani K., Ericsson L.H., Walsh K.A., Neurath H.;
RT "Amino-acid sequence of bovine carboxypeptidase B.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:1666-1670(1975).
RN [2]
RP SEQUENCE OF 31-93; 131-181; 263-265 AND 292-306.
RX MEDLINE=74260705; PubMed=4833744;
RA Schmid J.J., Hirs C.H.W.;
RT "Primary structure of bovine carboxypeptidase B. Inferences from the
RT locations of the half-cystines and identification of the active site
RT arginine.";
RL J. Biol. Chem. 249:3756-3764(1974).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND DISULFIDE BONDS.
RX MEDLINE=76265065; PubMed=957425;
RA Schmid M.P., Herrliott J.R.;
RT "Structure of carboxypeptidase B at 2.8-A resolution.";
RL J. Mol. Biol. 103:175-190(1976).
RN [4]
RP ACTIVE SITE.
RX MEDLINE=70007159; PubMed=5344132;
RA Plummer T.H. Jr.;
RT "Isolation and sequence of peptides at the active center of bovine
RT carboxypeptidase B.";
RL J. Biol. Chem. 244:5246-5253(1969).
RN [5]
RP ACTIVE SITE.
RX MEDLINE=73061487; PubMed=4565668;
RA Kimmel M.T., Plummer T.H. Jr.;
RT "Identification of a glutamic acid at the active center of bovine
RT carboxypeptidase B.";
RL J. Biol. Chem. 247:7864-7869(1972).
CC -!- CATALYTIC ACTIVITY: Peptidyl-L-lysine(or L-arginine) + H(2)O =
CC peptide + L-lysine(or L-arginine).
CC -!- SIMILARITY: Belongs to peptidase family M14.
DR PIR; A93797; CPB0B.
DR PDB; 1CPB; 30-SEP-83.
DR MEROPS; M14.003; -.
DR InterPro; IPR000834; Peptidase_M14.
DR Pfam; PF00246; Zn_carboxpept; 1.
DR PRINTS; PR00765; CRBOXYPTASEA.
DR SMART; SM00631; Zn_pept; 1.
DR PROSITE; PS00132; CARBOXYPEPT_ZN_1; 1.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.
KW Hydrolase; Carboxypeptidase; Metalloprotease; Zinc; 3D-structure.
FT DISULFID 63 76
FT DISULFID 135 158
FT DISULFID 149 163
FT METAL 66 66 ZINC.
FT METAL 69 69 ZINC.
FT METAL 194 194 ZINC.
FT ACT_SITE 246 246
FT ACT_SITE 268 268
FT ACT_SITE 306 AA; 34612 MW; C329D2655C4A172 CRC64;
SQ SEQUENCE 306 AA; 34612 MW; C329D2655C4A172 CRC64;

Query Match 46.3%; Score 38; DB 1; Length 306;
Best Local Similarity 35.7%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

QY 1 CAWVASRGIRPVGR 14  
Db 76 CQFVREAVRTYGR 89  
RESULT 21  
CYCR CHRVI  
ID -CYCR CHRVI STANDARD; PRT; 383 AA.  
AC O82947; 2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Photosynthetic reaction center cytochrome C subunit precursor.  
DE PUPC.  
GN Chromatium vinosum.  
OS Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;  
OC Chromatiaceae; Allochromatium.  
OX NCBI\_TaxID=1049;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=D / ATCC 17899 / DSM 180;  
RA Corson G.E., Nagashima K.V., Mateaura K., Sakuragi Y., Ruwanthi W.,  
RA Qin H., Allen R., Knaff D.B.;  
RT "Primary structure of genes encoding light-harvesting and reaction  
RT center proteins from Chromatium vinosum";  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A  
CC TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO  
CC OXIDIZED PRIMARY ELECTRON DONOR.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
CC (By similarity).  
CC -!- PTM: Binds 4 heme groups per molecule.  
CC -!- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER  
CC CYTOCHROME C SUBUNITS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; AB011811; BAA32742.1; -.  
DR HSSP; P07173; 6PRC.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR003158; CytC\_RC.  
DR Pfam; PF02276; CytoC\_RC; 1.  
DR PIRSF; PIRSF000017; RC cytochrome; 1.  
DR PRODom; PRO10011; CytC\_RC; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 4.  
DR Electron transport; Photosynthesis; Reaction center; Heme; Membrane;  
KW Lipoprotein; Repeat; Signal; Palmitate.  
FT SIGNAL 1 22 BY SIMILARITY.  
FT CHAIN 23 383 PHOTOSYNTHETIC REACTION CENTER CYTOCHROME  
FT C SUBUNIT.  
FT LIPID 23 23 S-diacylglycerol cysteine (By  
FT similarity).  
FT LIPID 23 23 N-palmitoyl cysteine (By similarity).  
FT METAL 94 94 IRON (HEME 1 AXIAL LIGAND)  
FT METAL 94 94 (BY SIMILARITY).  
FT BINDING 107 107 HEME 1 (COVALENT) (BY SIMILARITY).  
FT BINDING 110 110 HEME 2 (COVALENT) (BY SIMILARITY).  
FT METAL 111 111 IRON (HEME 1 AXIAL LIGAND)  
FT METAL 111 111 (BY SIMILARITY).  
FT METAL 130 130 IRON (HEME 2 AXIAL LIGAND)  
FT METAL 130 130 (BY SIMILARITY).  
FT METAL 144 144 IRON (HEME 4 AXIAL LIGAND)  
FT METAL 144 144 (BY SIMILARITY).  
FT BINDING 152 152 HEME 2 (COVALENT) (BY SIMILARITY).  
FT BINDING 155 155 HEME 2 (COVALENT) (BY SIMILARITY).  
FT METAL 156 156 IRON (HEME 2 AXIAL LIGAND)  
FT METAL 156 156 (BY SIMILARITY).

FT METAL 236 236 IRON (HEME 3 AXIAL LIGAND)  
FT BINDING 247 247 HEME 3 (COVALENT) (BY SIMILARITY).  
FT BINDING 250 250 HEME 3 (COVALENT) (BY SIMILARITY).  
FT METAL 251 251 IRON (HEME 3 AXIAL LIGAND)  
FT METAL 251 251 (BY SIMILARITY).  
FT BINDING 307 307 HEME 4 (COVALENT) (BY SIMILARITY).  
FT BINDING 310 310 HEME 4 (COVALENT) (BY SIMILARITY).  
FT METAL 311 311 IRON (HEME 4 AXIAL LIGAND)  
FT METAL 311 311 (BY SIMILARITY).  
SQ SEQUENCE 383 AA; 41522 MW; 96BCD91FF1B9AE7E CRC64;  
Query Match 46.3%; Score 38; DB 1; Length 383;  
Best Local Similarity 46.2%; Pred. No. 35;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
QY 2 AWVASRGIRPVGR 14  
Db 269 AWAIRHVRDINQ 281  
RESULT 22  
CBPB HUMAN  
ID CBPB HUMAN STANDARD; PRT; 417 AA.  
AC P15086; O60834; Q96B08;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Carboxypeptidase B precursor (EC 3.4.17.2) (Pancreas-specific protein)  
DE (PASP).  
DE CPB1 OR CPB OR PCPB.  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 16-40.  
RC TISSUE=Pancreas;  
RX MEDLINE=92129345; PubMed=1370825;  
RA Yamamoto K.K., Pousette A., Chow P., Wilson H., el Shami S.,  
RA French C.K.;  
RT "Isolation of a cDNA encoding a human serum marker for acute  
RT pancreatitis. Identification of pancreas-specific protein as  
RT pancreatic procarboxypeptidase B";  
RL J. Biol. Chem. 267:2575-2581(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RX MEDLINE=98182241; PubMed=9524066;  
RA Aloy P., Catasus L., Villegas V., Reverter D., Vendrell J.,  
RA Aviles F.X.;  
RT "Comparative analysis of the sequences and three-dimensional models  
RT of human procarboxypeptidases A1, A2 and B";  
RL Biol. Chem. 379:149-155(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Adrenal gland;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Xiaener R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Atschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Locuelli N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatane P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

```
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4].
RP SEQUENCE OF 16-43.
RC TISSUE=Pancreas;
RX MEDLINE=89153096; PubMed=2920728;
RA Pascual R., Burgos F.J., Soriano F., Mendez E., Aviles F.X.;
RT "Purification and properties of five different forms of human
RT procarboxypeptidases."
RN Eur. J. Biochem. 179:609-616(1989).
CC -!- CATALYTIC ACTIVITY: Pepidyl-L-lysine(or L-arginine) + H(2)O =
CC peptide + L-lysine(or L-arginine).
CC -!- TISSUE SPECIFICITY: Pancreas.
CC -!- SIMILARITY: Belongs to peptidase family M14.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M81057; AA866973.1; -.
CC EMBL; AJ224866; CAAL2163.1; -.
CC EMBL; BC015338; AAH5338.1; -.
CC PDB; 1KWM; 28-AUG-02.
CC MEROPS; M14.003; -.
CC Genew; HGNC:2299; CPB1.
CC MIM; 114852; -.
CC GO; GO:0004180; F:carboxypeptidase activity; TAS.
CC InterPro; IPR003146; Pept_M14A_propep.
CC InterPro; IPR000834; Peptidase_M14.
CC InterPro; IPR009020; Protease_inh1.
CC Pfam; PF02244; Propep_M14; 1.
CC Pfam; PF00246; Zn_Carboxept; 1.
CC PRINTS; PR00765; CRBOXYPTASEA.
CC SMART; SM00631; Zn_pept; 1.
CC PROSITE; PS00132; CARBOXYPEPT_ZN_1; 1.
CC PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.
CC Hydroxylase; Carboxypeptidase; Metalloprotease; Zinc; Zymogen; Signal;
KW 3D-structure.
FT SIGNAL 1 15
FT PROPEP 16 110 ACTIVATION PEPTIDE.
FT CHAIN 111 417 CARBOXYPEPTIDASE B.
FT METAL 176 176 ZINC (BY SIMILARITY).
FT METAL 179 179 ZINC (BY SIMILARITY).
FT METAL 304 304 ZINC (BY SIMILARITY).
FT ACT SITE 378 378 NUCLEOPHILE (BY SIMILARITY).
FT DISULFID 173 186 BY SIMILARITY.
FT DISULFID 245 268 BY SIMILARITY.
FT DISULFID 259 273 BY SIMILARITY.
FT CONFLICT 16 16 H -> A (IN REF. 1; AA SEQUENCE).
FT CONFLICT 17 17 H -> Q (IN REF. 1; AA SEQUENCE).
FT CONFLICT 37 37 H -> Q (IN REF. 4).
FT CONFLICT 208 208 D -> N (IN REF. 2).
FT CONFLICT 245 245 MISSING (IN REF. 1).
SQ SEQUENCE 417 AA; 47367 MW; E8BB98B27F5D5AF9 CRC64;

Query Match 46.3%; Score 38; DB 1; Length 417;
Best Local Similarity 35.7%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CANYASRGIRPVGR 14
Db 186 COMFVREAVTYGR 199

RESULT 23
ENO_MYCPU
```

```
ID ENO_MYCPU STANDARD; PRT; 456 AA.
AC Q98Q50;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN ENO OR MYPU_5180.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB C1P;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
RA Moszer I., Dybwig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis."
RL Nucleic Acids Res. 29:2145-2153(2001).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -!- COPACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the enolase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL445564; CAC13691.1; -.
CC PUR; F90576; F90576.
CC MyPulist; MYPU_5180; -.
CC HAVAP; MF_00318; -.
CC InterPro; IPR000941; Enolase.
CC Pfam; PF00113; enolase; 1.
CC Pfam; PF03952; enolase_N; 1.
CC PRINTS; PR00148; ENOLASE.
CC PRODOM; PD000902; Enolase; 1.
CC TIGRFAWS; TIGR01060; eno; 1.
CC PROSITE; PS00164; ENOLASE; 1.
CC Lyase; Glycolysis; Magnesium; Complete proteome.
KW ACT SITE 159 159 BY SIMILARITY.
FT ACT SITE 159 159 MAGNESIUM (BY SIMILARITY).
FT METAL 250 250 MAGNESIUM (BY SIMILARITY).
FT METAL 312 312 MAGNESIUM (BY SIMILARITY).
FT METAL 339 339 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 456 AA; 49877 MW; 330300BF0FDEA2 CRC64;

Query Match 46.3%; Score 38; DB 1; Length 456;
Best Local Similarity 62.5%; Pred. No. 41;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WYASRGIR 10
Db 62 WYASKGVQ 69

RESULT 24
SM6C_HUMAN STANDARD; PRT; 930 AA.
ID SM6C_HUMAN
AC Q9H3T2; Q8WXT8; Q8WXT9; Q8WXT0; Q96JF8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Semaphorin 6C precursor (Semaphorin Y) (Sema Y).
```

GN SEMA6C OR SEMAY OR KIAA1869.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Brain;  
RA Kimura T., Ishida H.;  
RN Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
[2]  
RN SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RC TISSUE=Brain;  
RA Qu X., Zhai Y., Wei H., Zhang C., King G., Yu Y., Wu S., Zhang Y.,  
RN Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
[3]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Brain;  
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;  
RN MEDLINE=21245130; PubMed=11347906;  
RA "Prediction of the coding sequences of unidentified human genes. XX.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 8:85-95(2001).  
CC -!- FUNCTION: May be a stop signal for the dorsal root ganglion  
CC neurons in their target areas, and possibly also for other  
CC neurons. May also be involved in the maintenance and remodeling of  
CC neuronal connections (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=1; Synonyms=Short 1;  
CC IsoId=Q9H3T2-1; Sequence=Displayed;  
CC Name=2; Synonyms=Short 2;  
CC IsoId=Q9H3T2-2; Sequence=VSP\_006046; VSP\_006047;  
CC Name=3; Synonyms=Long;  
CC IsoId=Q9H3T2-3; Sequence=VSP\_006047;  
CC -!- SIMILARITY: Belongs to the semaphorin family.  
CC -!- SIMILARITY: Contains 1 Sema domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AB022434; BAB20670.1; -  
CC EMBL; AF339152; AAL72098.1; -  
CC EMBL; AF339153; AAL72099.1; -  
CC EMBL; AF339154; AAL72100.1; -  
CC EMBL; AB058772; BAB47498.1; ALT\_INIT.  
CC Genew; HGNC:10740; SEMA6C.  
CC InterPro; IPR003659; Plexin-like.  
CC InterPro; IPR001627; Sema.  
CC Pfam; PF01403; Sema; 1.  
CC SMART; SM00423; PSI; 1.  
CC SMART; SM00630; Sema; 1.  
CC Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;  
CC Developmental protein; Alternative splicing.  
FT SIGNAL 1 24  
FT CHAIN 25 930  
FT DOMAIN 25 604  
FT TRANSMEM 605 625  
FT DOMAIN 626 930  
FT DOMAIN 233 540  
FT DOMAIN 662 667  
FT DOMAIN 752 755  
FT CARBOHYD 70 70  
FT CARBOHYD 286 286  
FT CARBOHYD 437 437

FT VARSPLIC 184 223 Missing (in isoform 2).  
FT VARSPLIC 586 586 /FTid=VSP\_006046.  
FT Y -> YVLPGPSPGTPSPSPSDAHPRPQSSTLGVHTR  
FT (in isoform 2 and isoform 3).  
FT /FTid=VSP\_006047.  
FT I -> V (IN REF. 2; AAL72099).  
FT CONFLICT 125 125  
FT CONFLICT 252 252  
FT CONFLICT 455 455 R -> K (IN REF. 1).  
FT CONFLICT 455 455 P -> T (IN REF. 1 AND 3).  
SQ SEQUENCE 930 AA; 99682 MW; 8AF8814ADBC94C88 CRC64;  
Query Match 46.3%; Score 38; DB 1; Length 930;  
Best Local Similarity 62.5%; Pred. No. 82;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CAWVASRG 8  
Db 537 CGWHSRG 544  
RESULT 25  
BCAT CAEEL  
ID BCAT CAEEL STANDARD; PRT; 415 AA.  
AC P54688;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUL-1998 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Branched-chain amino acid aminotransferase, cytosolic (EC 2.6.1.42)  
DE (BCAT) (ECA39 protein).  
OS ECA-39 OR ECA39 OR K02A4.1.  
GN Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA McMurray A.A.;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 21-398 FROM N.A.  
RX MEDLINE=96293490; PubMed=8692959;  
RA Schuldiner O., Eden A., Ben-Yosef T., Yanuka O., Simchen G.,  
RA Benvenisty N.;  
RT "ECA39, a conserved gene regulated by c-Myc in mice, is involved in  
RT G1/S cell cycle regulation in yeast";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:7143-7148(1996).  
CC -!- FUNCTION: Catalyzes the first reaction in the catabolism of the  
CC essential branched chain amino acids leucine, isoleucine, and  
CC valine.  
CC -!- CATALYTIC ACTIVITY: L-leucine + 2-oxoglutarate = 4-methyl-2-  
CC oxopentanoate + L-glutamate.  
CC -!- COFACTOR: Pyridoxal phosphate.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to class-IV of pyridoxal-phosphate-dependent  
CC aminotransferases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; 267883; CAA91805.1; -  
CC EMBL; U21550; AAC47236.1; -  
CC PIR; T23215; T23215.  
CC HSP; P00510; IASG.  
CC WormPep; K02A4.1; C03457.  
CC InterPro; IPR001544; Aminotrans IV.  
CC InterPro; IPR005786; B.amino.transII.  
CC Pfam; PF01063; aminotran\_4; 1.  
CC ProDom; PD001961; Aminotran\_4; 1.



```
DR TIGRFAMs; TIGR01123; ilvE II; 1.
DR PROSITE; PS00770; AA TRANSFER CLASS 4; 1.
KW Transferrase; Amino transferase; Branched-chain amino acid biosynthesis;
FT BINDING 244 244 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT CONFLICT 86 86 D -> I (IN REF. 2).
FT CONFLICT 118 118 F -> FS (IN REF. 2).
FT CONFLICT 154 170 DFDSSEMINVLTELLRL -> VFRLRGDDQCTYRILRS
      (IN REF. 2).
FT CONFLICT 391 397 MCKFYNT -> RKILQHN (IN REF. 2).
SQ SEQUENCE 415 AA; 47341 MW; AE216PC9623PB390 CRC64;
Query Match 45.7%; Score 37.5; DB 1; Length 415;
Best Local Similarity 44.4%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 4; Indels 5; Gaps 1;

QY 1 CAWYASRG-----IRPVG 13
DB 83 CWDADRGWHHPKIEPIG 100

RESULT 26
R157 BOVIN STANDARD; PRT; 147 AA.
AC Q28183;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Retina-specific 15.7 kDa protein.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=86258168; PubMed=2425311;
RA Nakagawa Y., Kuo C.H., Ishii K., Shiosaka S., Tohyama M., Miki N.;
RT "Cloning and characterization of a cDNA specific for bovine retina.";
RL Neurosci. Res. 3:300-310(1996).
CC -!- TISSUE SPECIFICITY: Retina.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; M34915; AAA30756.1; -.
DR F019; I45978; I45978.
SQ SEQUENCE 147 AA; 15658 MW; 1FEDA48788B3645 CRC64;

Query Match 45.1%; Score 37; DB 1; Length 147;
Best Local Similarity 71.4%; Pred. No. 20;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAWYASR 7
DB 100 CSWFASR 106

RESULT 27
Y064 TREPA STANDARD; PRT; 191 AA.
AC O83103;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0064.
GN TP0064.

Treponema pallidum.
Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khaalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Attiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT Spirochete.";
RL Science 281:375-388(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; AE001191; AAC65064.1; -.
DR F019; H71370; H71370.
DR TIGR; TP0064; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 191 AA; 22079 MW; 2837931F910CD831 CRC64;

Query Match 45.1%; Score 37; DB 1; Length 191;
Best Local Similarity 57.1%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CAWYASRGIRPVGR 14
DB 16 CAPYGRGGEFVR 29

RESULT 28
DAPF CHLTE STANDARD; PRT; 257 AA.
AC Q8KAX9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Diaminopimelate epimerase (EC 5.1.1.7) (DAP epimerase).
GN DAPF OR CT2021.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwin M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC -!- CATALYTIC ACTIVITY: L,L-2,6-diaminoheptanedioate = meso-
CC diaminohexanedioate.
CC -!- PATHWAY: Biosynthesis of lysine from aspartate semialdehyde; sixth
CC step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
```

```

CC CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U32734; AAC22182.1; -.
CC PIR; C64074; C64074.
CC HSSP; P11604; IBS7.
DR TIGR; H10524; -.
DR InterPro; IPRO06411; Fruct_bisp_bact.
DR InterPro; IPRO00771; K_bp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; K_bp_aldolase; 1.
DR TIGRFAMs; TIGR00167; cbbA; 1.
DR TIGRFAMs; TIGR01520; FruBisaldo_II_A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR InterPro; IPS00806; ALDOLASE_CLASS_II_2; 1.
DR Lyase; Glycolysis; Zinc; Complete proteome.
KW METAL 108 ZINC (BY SIMILARITY).
FT METAL 111 111 ZINC (BY SIMILARITY).
SQ SEQUENCE 359 AA; 33339 MW; 1EDDFCDD0B69E32C CRC64;
Query Match 45.1%; Score 37; DB 1; Length 359;
Best Local Similarity 60.0%; Pred. No. 49;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 AWYASRGIRP 11
DB 67 AFYAGKGIKP 76
-----
RESULT 30
ALR1_RHILO
ID ALR1_RHILO STANDARD; PRT; 397 AA.
AC Q98A05;
CD 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alanine racemase, biosynthetic (EC 5.1.1.1).
GN ALR OR ML6211.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Igesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RT DNA Res. 7:331-338(2000).
RL -!- FUNCTION: Provides the D-alanine required for cell wall
CC biosynthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: L-alanine = D-alanine.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; first
CC step.
CC -!- SIMILARITY: Belongs to the alanine racemase family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; AP003008; BAB52539.1; -.
DR HAMAP; MF_01201; ?; 1.
DR InterPro; IPRO00821; Ala racemase.

```

```

DR InterPro: IPR000906; Racem decarbox_C.
DR InterPro: IPR001508; UPR0001.
DR Pfam: PF00842; Ala_racemase_C; 1.
DR Pfam: PF01168; Ala_racemase_N; 1.
DR PRINTS: PR00992; ALARACEMASE.
DR PROSITE: PS00395; ALANINE RACEMASE; 1.
KW Isomerase; Pyridoxal phosphate; Cell wall; Peptidoglycan synthesis;
KW Complete proteome.
FT ACT_SITE 42 42 CATALYTIC BASE SPECIFIC TO D-ALANINE
(BY SIMILARITY).
FT ACT_SITE 257 257 CATALYTIC BASE SPECIFIC TO L-ALANINE
(BY SIMILARITY).
FT BINDING 42 42 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT SEQUENCE 397 AA; 43462 MW; 15996F5136A9C307 CRC64;
SQ
Query Match 45.1%; Score 37; DB 1; Length 397;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 WYASRGIRPV 12
|||:|:|
Db 98 WYRSHGLIPV 107

RESULT 31
SR54 HALN1 STANDARD; PRT; 460 AA.
AC Q9HWN5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Signal recognition 54 kDa protein (SRP54).
GN SRP54 OR VNG2459G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebbardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000)
CC -!- FUNCTION: Binds to the signal sequence of presecretory protein
when they emerge from the ribosomes (By similarity).
CC -!- SUBUNIT: Archaeal signal recognition particle consists of a 7S RNA
molecule of 300 nucleotides and two protein subunits: SRP54 and
SRP19 (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- DOMAIN: Has a two domain structure: the G-domain binds GTP; the M-
domain binds the 7S RNA in presence of SRP19 and also binds the
signal sequence (By similarity).
CC -!- SIMILARITY: Belongs to the GTP-binding SRP family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AE005125; AAG20536.1; -.
CC F01; D84396; D84396.
CC HSSP; O07347; 2FFH.
CC HAMAP; MF_00306; -.

```

```

DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR000897; SRP54.
DR InterPro: IPR004135; SRP54_SPB.
DR Pfam: PF00448; SRP54; 1.
DR Pfam: PF02881; SRP54_N; 1.
DR Pfam: PF02978; SRP_SPB; 1.
DR ProDom: PD000819; SRP54; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00300; SRP54; FALSE NEG.
KW Signal recognition particle; GTP-binding; RNA-binding;
KW Complete proteome.
FT DOMAIN 1 289 G-DOMAIN (BY SIMILARITY).
FT DOMAIN 290 460 M-DOMAIN (BY SIMILARITY).
FT NP_BIND 104 111 GTP (BY SIMILARITY).
FT NP_BIND 184 188 GTP (BY SIMILARITY).
FT NP_BIND 242 245 GTP (BY SIMILARITY).
FT DOMAIN 449 458 POLY-GLY.
FT SEQUENCE 460 AA; 50149 MW; 8361B782E651352E CRC64;
SQ
Query Match 45.1%; Score 37; DB 1; Length 460;
Best Local Similarity 44.4%; Pred. No. 62;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 WYASRGIRP 11
|||:|:|
Db 120 WFSKGLR 128

RESULT 32
SR54 HALVO STANDARD; PRT; 465 AA.
AC Q977V2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Signal recognition 54 kDa protein (SRP54).
GN SRP54.
OS Halobacterium volcanii (Haloferax volcanii).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloferax.
OX NCBI_TaxID=2246;
RN [1]
RP SEQUENCE FROM N.A.
RA Rose R.W., Pohlschroder M.;
Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds to the signal sequence of presecretory protein
when they emerge from the ribosomes (By similarity).
CC -!- SUBUNIT: Archaeal signal recognition particle consists of a 7S RNA
molecule of 300 nucleotides and two protein subunits: SRP54 and
SRP19 (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- DOMAIN: Has a two domain structure: the G-domain binds GTP; the M-
domain binds the 7S RNA in presence of SRP19 and also binds the
signal sequence (By similarity).
CC -!- SIMILARITY: Belongs to the GTP-binding SRP family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF395887; AAK93963.1; -.
CC HAMAP; MF_00306; -.
CC InterPro: IPR003593; AAA_ATPase.
CC InterPro: IPR000897; SRP54.
CC InterPro: IPR004125; SRP54_SPB.
CC Pfam; PF00448; SRP54; 1.
CC Pfam; PF02881; SRP54_N; 1.
CC Pfam; PF02978; SRP_SPB; 1.
CC ProDom; PD000819; SRP54; 1.

```

```

DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00300; SRP54; FALSE NEG.
KW Signal recognition particle; GTP-binding; RNA-binding.
FT DOMAIN 1 289 G-DOMAIN (BY SIMILARITY).
FT DOMAIN 290 465 M-DOMAIN (BY SIMILARITY).
FT NP_BIND 104 111 GTP (BY SIMILARITY).
FT NP_BIND 184 188 GTP (BY SIMILARITY).
FT NP_BIND 242 245 GTP (BY SIMILARITY).
SQ SEQUENCE 465 AA; 50918 MW; 9F5705838DFF8370 CRC64;

Query Match 45.1%; Score 37; DB 1; Length 465;
Best Local Similarity 44.4%; Pred. No. 62;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 3 WYASRGIRP 11
Db 120 WFSKKGLRP 128

RESULT 33
NCAP CCHVF STANDARD; PRT; 482 AA.
AC P27317;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Nucleocapsid protein (Nucleoprotein).
GN N.
OS Crimean-Congo hemorrhagic fever virus (isolate C68031) (CCHFV).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Nairovirus.
OX NCBI_TaxID=11594;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92351584; PubMed=1641991;
RA Marriott A.C.; Nuttall P.A.;
RT "Comparison of the S RNA segments and nucleoprotein sequences of
RL Crimean-Congo hemorrhagic fever, Hazara, and Dugbe viruses.";
RL Virology 189:795-799 (1992).
CC -!- SUBCELLULAR LOCATION: Internal protein of virus particle.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M86625; AAA42888.1; -
DR PIR; B42990; VHVUCH.
DR InterPro; IPR003486; Nairo nucleocap.
DR Pfam; PF02477; Nairo nucleocap; 1.
DR ProDom; PD006459; Nairo nucleocap; 1.
KW Nucleocapsid, RNA-binding.
SQ SEQUENCE 482 AA; 53965 MW; 171F8D9438F00FA2 CRC64;

Query Match 45.1%; Score 37; DB 1; Length 482;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAWVASRGI 9
Db 80 CAWVSSTGI 88

RESULT 34
ID BOP1 HUMAN STANDARD; PRT; 746 AA.
AC Q14137; Q96926; Q961S6; Q9BSA7; Q9BVN0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Block of proliferation 1.

```

```

GN BOP1 OR KIAA0124.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, Eye, Lymph, Muscle, and Pancreas;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Altshuler S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.B.; Scheetz T.B.;
RA Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA Raha S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.;
RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
RA Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA Fahey J.; Helton E.; Kettman M.; Madan A.; Rodriguez S.; Sanchez A.;
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.;
RA Butterfield Y.S.N.; Krzywinski M.I.; Skalska U.; Smalls D.E.;
RA Schnerch A.; Schein J.E.; Jones S.J.M.; Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE OF 65-746 FROM N.A.
RX MEDLINE=96127530; PubMed=8590280;
RA Nagase T.; Seki N.; Tanaka A.; Ishikawa K.-I.; Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. IV.
RT The coding sequences of 40 new genes (K1AA0121-K1AA0160) deduced by
RL DNA Res. 2:167-174 (1995).
RN [3]
RP SUBCELLULAR LOCATION.
RX MEDLINE=22312777; PubMed=12429849;
RA Scherl A.; Coute Y.; Deon C.; Calle A.; Kindbeiter K.; Sanchez J.-C.;
RA Greco A.; Hochstrasser D.; Diaz J.J.;
RT "Functional proteomic analysis of human nucleolus.";
RL Mol. Biol. Cell 13:4100-4109 (2002).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -!- SIMILARITY: Contains 7 WD repeats.
CC -!- SIMILARITY: STRONG, TO YEAST YMR049C.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; BC001086; AAH01086.1; ALT_INIT.
DR EMBL; BC005160; AAH05160.1; ALT_INIT.
DR EMBL; BC007274; AAH07274.1; -
DR EMBL; BC013787; AAH13787.1; -
DR EMBL; BC013980; AAH13980.1; -
DR EMBL; BC017674; AAH17674.1; -
DR EMBL; D50914; BAA09473.1; -
DR SWISS-2DPAGE; Q14137; HUMAN.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 6.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 2.
DR PROSITE; PS50294; WD_REPEATS_REGION; 2.
KW Nuclear protein; Repeat; WD repeat.

```

```

FT REPEAT 411 450 WD 1.
FT REPEAT 452 492 WD 2.
FT REPEAT 532 576 WD 3.
FT REPEAT 577 615 WD 4.
FT REPEAT 618 657 WD 5.
FT REPEAT 681 700 WD 6.
FT REPEAT 716 745 WD 7.
FT REPEAT 777 777 R -> H (IN REF. 2).
SQ SEQUENCE 746 AA; 83629 MW; 49A97BE21B0E3DD CRC64;

Query Match 45.1%; Score 37; DB 1; Length 746;
Best Local Similarity 38.5%; Pred. No. 99;
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 CAWYASRGIRPVG 13
DB 620 CKWSSSLAVHPAG 632

RESULT 35
ID METE_LACPL STANDARD; PRT; 768 AA.
AC Q88X63;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 5-methyltetrahydropteroyltri-L-glutamate--homocysteine methyltransferase
DE (EC 2.1.1.14) (Methionine synthase, vitamin-B12 independent isozyme)
DE (Cobalamin-independent methionine synthase).
METE OR LP 1375.
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCIMB 8826 / WCP81;
RX MEDLINE=22480296; PubMed=12566566;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.B.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCP81.";
RL Proc. Natl Acad. Sci. U.S.A. 100:1990-1995(2003).
CC -!- FUNCTION: Catalyzes the transfer of a methyl group from 5-
CC methyltetrahydrofolate to homocysteine resulting in methionine
CC formation (By similarity).
CC -!- CATALYTIC ACTIVITY: 5-methyltetrahydropteroyltri-L-glutamate + L-
CC homocysteine = tetrahydropteroyltri-L-glutamate + L-methionine.
CC -!- COFACTOR: Zinc; binds one ion per subunit (By similarity).
CC -!- PATHWAY: Terminal step in the de novo biosynthesis of methionine.
CC -!- SIMILARITY: Belongs to the vitamin-B12 independent methionine
CC synthase family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AL935256; CAD63852.1; --
DR HAMAP; MF 00172; -- 1.
DR InterPro; IPR002629; Methionine synth.
DR Pfam; PF0117; Methionine_synth; 1.
DR Transferrase; Methyltransferase; Methionine biosynthesis; Zinc; Repeat;
KW Complete proteome.
FT METAL 650 650 ZINC (BY SIMILARITY).
FT METAL 652 652 ZINC (BY SIMILARITY).
FT METAL 735 735 ZINC (BY SIMILARITY).
FT SEQUENCE 768 AA; 86432 MW; 477E4CB0C87ECE24 CRC64;

```

```

Query Match 45.1%; Score 37; DB 1; Length 768;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 YASRGIRP 11
DB 521 YGSRGVRP 528

RESULT 36
ID EPB4_MOUSE STANDARD; PRT; 987 AA.
AC P54761; Q60627;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ephrin type-B receptor 4 precursor (EC 2.7.1.112) (Tyrosine-protein
DE kinase receptor MDK-2) (Developmental kinase 2) (Tyrosine kinase MYK-
DE 1).
GN EPB4 OR MDK2 OR MYK1 OR HTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/c; TISSUE=Kidney;
RX MEDLINE=96074837; PubMed=7479528;
RA Ciossek T., Lerch M.M., Ullrich A.;
RT "Cloning, characterization, and differential expression of MDK2 and
RT MDK5, two novel receptor tyrosine kinases of the eck/eph family.";
RL Oncogene 11:2085-2095(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/c; TISSUE=Lung;
RX MEDLINE=94203677; PubMed=8152808;
RA Andres A.C., Reid H.H., Zurcher G., Blaschke R.J., Albrecht D.,
RA Ziemiński A.;
RT "Expression of two novel eph-related receptor protein tyrosine
RT kinases in mammary gland development and carcinogenesis.";
RL Oncogene 9:1461-1467(1994).
CC -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY. BINDS TO
CC EPHRIN-B2.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: DEVELOPING CARDIOVASCULAR SYSTEM.
CC -!- DEVELOPMENTAL STAGE: ABUNDANT EXPRESSION AT E16.5 IN VARIOUS ORGAN
CC SYSTEMS, INCLUDING THYMUS, HEART, LUNGS AND KIDNEY, WHICH APPEARED
CC TO BE ASSOCIATED WITH CELLS OF ENDOTHELIAL ORIGIN.
CC -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Ephrin
CC receptor subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Z49085; CAA88909.1; --
DR EMBL; U06834; AAA18591.1; --
DR PIR; I48552; I48652.
DR PIR; I48953; I48953.
DR HSP; P29323; IB4F.
DR MGD; MG1:104757; Edb4.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR001090; Ephrin receptor.
DR InterPro; IPR008957; FN_III-like.

```

DR InterPro; IPR003961; FN III.  
 DR InterPro; IPR003962; FNIII subd.  
 DR InterPro; IPR008979; Gal bind like.  
 DR InterPro; IPR000719; ProT\_kinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase AS.  
 DR InterPro; IPR001426; YKase\_receptorV.  
 DR Pfam; PF01404; EPH\_lbd; 1.  
 DR Pfam; PF00041; fn3\_2.  
 DR Pfam; PF00069; Ekinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00014; FNTYPEIII.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD001495; Ephrin\_receptor; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00615; EPH\_lbd; 1.  
 DR SMART; SM00650; FN3; 2.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS01186; EGF\_2; UNKNOWN 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 DR PROSITE; PS00790; RECEPTOR\_TYR\_KIN\_V1; 1.  
 DR PROSITE; PS00791; RECEPTOR\_TYR\_KIN\_V2; 1.  
 DR PROSITE; PS00103; SAM\_DOMAIN; 1.  
 DR Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat.  
 FT SIGNAL 1 15  
 FT CHAIN 16 987  
 FT DOMAIN 16 539  
 FT TRANSMEM 540 987  
 FT DOMAIN 561 987  
 FT DOMAIN 184 320  
 FT DOMAIN 321 429  
 FT DOMAIN 430 528  
 FT DOMAIN 515 899  
 FT DOMAIN 907 971  
 FT SITE 985 987  
 FT NP\_BIND 621 629  
 FT BINDING 647 647  
 FT ACT\_SITE 740 740  
 FT MOD\_RES 590 590  
 FT MOD\_RES 596 596  
 FT MOD\_RES 774 774  
 FT MOD\_RES 924 924  
 FT CARBOHYD 203 203  
 FT CARBOHYD 335 335  
 FT CARBOHYD 426 426  
 FT CONFLICT 351 353  
 FT CONFLICT 389 389  
 FT CONFLICT 659 659  
 FT CONFLICT 783 783  
 FT CONFLICT 805 805  
 FT CONFLICT 913 913  
 FT CONFLICT 938 938  
 FT CONFLICT 970 971  
 FT SEQUENCE 987 AA; 108922 MW; 471F809579D98A02 CRC64;  
 Query Match 45.1%; Score 37; DB 1; Length 987;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 WYASRGIRP 11  
 DB 390 WVAIRGIRP 398  
 RESULT 37  
 Y4VI\_RHISN  
 ID\_Y4VI\_RHISN STANDARD; PRT; 548 AA.  
 AC Q53217;  
 InterPro; IPR003961; FN III.  
 InterPro; IPR003962; FNIII subd.  
 InterPro; IPR008979; Gal bind like.  
 InterPro; IPR000719; ProT\_kinase.  
 InterPro; IPR001660; SAM.  
 InterPro; IPR001245; Tyr\_kinase.  
 InterPro; IPR008266; Tyr\_kinase AS.  
 InterPro; IPR001426; YKase\_receptorV.  
 Pfam; PF01404; EPH\_lbd; 1.  
 Pfam; PF00041; fn3\_2.  
 Pfam; PF00069; Ekinase; 1.  
 Pfam; PF00536; SAM; 1.  
 PRINTS; PR00014; FNTYPEIII.  
 PRINTS; PR00109; TYRKINASE.  
 ProDom; PD001495; Ephrin\_receptor; 1.  
 ProDom; PD000001; Prot\_kinase; 1.  
 SMART; SM00615; EPH\_lbd; 1.  
 SMART; SM00650; FN3; 2.  
 SMART; SM00454; SAM; 1.  
 SMART; SM00219; Tyrc; 1.  
 PROSITE; PS01186; EGF\_2; UNKNOWN 1.  
 PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 PROSITE; PS00790; RECEPTOR\_TYR\_KIN\_V1; 1.  
 PROSITE; PS00791; RECEPTOR\_TYR\_KIN\_V2; 1.  
 PROSITE; PS00103; SAM\_DOMAIN; 1.  
 Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat.  
 FT SIGNAL 1 15  
 FT CHAIN 16 987  
 FT DOMAIN 16 539  
 FT TRANSMEM 540 987  
 FT DOMAIN 561 987  
 FT DOMAIN 184 320  
 FT DOMAIN 321 429  
 FT DOMAIN 430 528  
 FT DOMAIN 515 899  
 FT DOMAIN 907 971  
 FT SITE 985 987  
 FT NP\_BIND 621 629  
 FT BINDING 647 647  
 FT ACT\_SITE 740 740  
 FT MOD\_RES 590 590  
 FT MOD\_RES 596 596  
 FT MOD\_RES 774 774  
 FT MOD\_RES 924 924  
 FT CARBOHYD 203 203  
 FT CARBOHYD 335 335  
 FT CARBOHYD 426 426  
 FT CONFLICT 351 353  
 FT CONFLICT 389 389  
 FT CONFLICT 659 659  
 FT CONFLICT 783 783  
 FT CONFLICT 805 805  
 FT CONFLICT 913 913  
 FT CONFLICT 938 938  
 FT CONFLICT 970 971  
 FT SEQUENCE 987 AA; 108922 MW; 471F809579D98A02 CRC64;  
 Query Match 45.1%; Score 37; DB 1; Length 987;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 WYASRGIRP 11  
 DB 390 WVAIRGIRP 398  
 RESULT 37  
 Y4VI\_RHISN  
 ID\_Y4VI\_RHISN STANDARD; PRT; 548 AA.  
 AC Q53217;

DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Putative short-chain type dehydrogenase/reductase Y4VI (EC 1.-.-.-).  
 GN Y4VI.  
 OS Rhizobium sp. (strain NGR234).  
 OG Plasmid sym pNGR234a.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
 OC NCBI\_TaxID=394;  
 RN [1]\_TaxID=394;  
 RP SEQUENCE FROM N.A. PubMed=9163424;  
 RX MEDLINE=97305956; PubMed=9163424;  
 RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,  
 RA Perret X.;  
 RT "Molecular basis of symbiosis between Rhizobium and legumes.";  
 RL Nature 387:394-401(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. PubMed=8796346;  
 RX MEDLINE=96389014; PubMed=8796346;  
 RA Freiberg C., Perret X., Broughton W.J., Rosenthal A.;  
 RT "Sequencing the 500-kb GC-rich symbiotic replicon of Rhizobium sp.  
 NGR234 using dye terminators and a thermostable 'sequenase'; a  
 beginning.";  
 RL Genome Res. 6:590-600(1996).  
 CC -! SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 (SDR) FAMILY. CONTAINS TWO DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; Z68203; CAA92424.1; ALT INIT.  
 CC EMBL; AE000101; AAB91897.1; -  
 CC HSSP; P50163; 2AE1.  
 CC InterPro; IPR002198; ADH\_short.  
 CC Pfam; PF00106; adh\_short; 2.  
 CC PRINTS; PR00080; SDRFAMILY.  
 CC PROSITE; PS00061; ADH\_SHORT; 1.  
 CC KW Hypothetical protein; Oxidoreductase; Plasmid; Repeat;  
 FT DOMAIN 1 250  
 FT DOMAIN 271 548  
 FT NP\_BIND 12 37  
 FT ACT\_SITE 154 154  
 FT NP\_BIND 280 304  
 FT ACT\_SITE 420 420  
 FT SEQUENCE 548 AA; 57593 MW; 399598C5DF17C23F CRC64;  
 Query Match 44.5%; Score 36.5; DB 1; Length 548;  
 Best Local Similarity 80.0%; Pred. No. 90;  
 Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 CAWVASRGIR 10  
 DB 170 CRW-ASRGIR 178  
 RESULT 38  
 ENRN\_BPT7  
 ID ENRN\_BPT7 STANDARD; PRT; 149 AA.  
 AC P00641;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Endodeoxyribonuclease I (EC 3.1.21.2) (endonuclease).  
 GN 3.  
 OS Bacteriophage T7.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;

```

OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83241725; PubMed=6864790;
RA Dunn J.J., Studier F.W.;
RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
RT locations of T7 genetic elements.";
RL J. Mol. Biol. 166:477-535(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=82078034; PubMed=7310871;
RA Dunn J.J., Studier F.W.;
RT "Nucleotide sequence from the genetic left end of bacteriophage T7
RT DNA to the beginning of gene 4.";
RL J. Mol. Biol. 148:303-330(1981).
CC -!- FUNCTION: ENDOXYRIBONUCLEASE I, WHICH IS EXPRESSED IN THE LATE
CC STAGE, IS NECESSARY FOR T7 GENETIC RECOMBINATION AND THE BREAKDOWN
CC OF HOST DNA. IN THE EARLY STAGE OF INFECTION, T7 DNA REPLICATES AS
CC A LINEAR MONOMER. IN THE LATE STAGE, THE T7 DNA REPLICATES VIA
CC LINEAR CONCATAMERS SEVERAL GENOMES IN LENGTH. THE GENE 3 PRODUCT
CC HAS ALSO BEEN IMPLICATED IN THE MATURATION OF THESE CONCATAMERS.
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphooligonucleotide end-products.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01127; CAA24345.1; -
CC EMBL; V01146; CAA24402.1; -
CC PIR; B94615; NEBP37.
CC PDB; 1FZR; 10-OCT-01.
CC PDB; 1MOD; 10-JUL-02.
CC PDB; 1MOI; 18-DEC-02.
CC InterPro; IPR008029; Phage endo I.
CC Pfam; PF05367; Phage endo I_1.
CC Hydrolase; Nuclease; Endonuclease; 3D-structure.
SQ SEQUENCE 149 AA; 17172 MW; D092AA28E3743BC1 CRC64;

Query Match 43.9%; Score 36; DB 1; Length 149;
Best Local Similarity 58.3%; Pred. No. 31;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVG 13
Db 2 AGYGAKGIRKVG 13

RESULT 39
TIM2_RABIT
ID TIM2_RABIT STANDARD; PRT; 194 AA.
AC Q9TEZ7; O97589;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Metalloprotease inhibitor 2 (TIMP-2) (Tissue inhibitor of
DE metalloproteinases-2).
GN Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96049520; PubMed=8548358;
RA Wertheimer S.J., Katz S.L.;
RT "Molecular cloning and characterization of rabbit TIMP2.";
RT Inflamm. Res. 44:S121-S122(1995).

```

```

[2]
RN SEQUENCE OF 17-154 FROM N.A.
RP STRAIN=New Zealand white;
RX MEDLINE=99057575; PubMed=9837780;
RA Reno C., Boykiw R., Martinez M.L., Hart D.A.;
RT "Temporal alterations in mRNA levels for proteinases and inhibitors
RT and their potential regulators in the healing medial collateral
RT ligament.";
RL Biochem. Biophys. Res. Commun. 252:757-763(1998).
CC -!- FUNCTION: Complexes with metalloproteinases (such as collagenases)
CC and irreversibly inactivates them.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: The activity of TIMP2 is dependent on the presence of
CC disulfide bonds (by similarity).
CC -!- SIMILARITY: Belongs to the TIMP family.
CC -!- SIMILARITY: Contains 1 NTR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF069713; AAC95005.1; -
CC HSP; P16035; IBR9.
CC InterPro; IPR001820; TIMP.
CC InterPro; IPR008993; TIMP_like.
CC Pfam; PF00965; TIMP; 1.
CC SMART; SMC0206; NTR; 1.
CC PROSITE; PS00189; NTR; 1.
CC PROSITE; PS00288; TIMP; 1.
CC Metalloprotease inhibitor.
FT DOMAIN 1 126
FT DISULFID 1 172
FT DISULFID 3 101
FT DISULFID 13 126
FT DISULFID 128 175
FT DISULFID 133 138
FT DISULFID 146 167
FT DISULFID 17 17
FT CONFLICT 25 26
FT CONFLICT 58 58
FT CONFLICT 78 78
FT CONFLICT 93 95
FT CONFLICT 109 109
FT CONFLICT 112 112
FT CONFLICT 131 131
SQ SEQUENCE 194 AA; 21849 MW; CDC810A2D38C4A9 CRC64;

Query Match 43.9%; Score 36; DB 1; Length 194;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 1 CAWYASRGIRP 11
Db 175 CAWY--RGAAP 183

RESULT 40
TIM2_CRIL0
ID TIM2_CRIL0 STANDARD; PRT; 196 AA.
AC Q60453;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Metalloprotease inhibitor 2 precursor (TIMP-2) (Tissue inhibitor of
DE metalloproteinases-2) (Fragment).
GN TIMP2.
OC Cricetus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

```

```

OC Cricetulus.
OX NCBI_TaxID=10030;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Suzuki Y.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Complexes with metalloproteinases (such as collagenases)
CC and irreversibly inactivates them (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: The activity of TIMP2 is dependent on the presence of
CC disulfide bonds (By similarity).
CC -!- SIMILARITY: Belongs to the TIMP family.
CC -!- SIMILARITY: Contains 1 NTR domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75924; CAA53528.1; -.
CC PIR; S38624; S38624.
CC HSSP; P16035; 1BR9.
CC InterPro; IPR001820; TIMP.
CC InterPro; IPR008993; TIMP_like.
CC Pfam; PF00965; TIMP; 1.
CC SMART; SM00206; NTR; 1.
CC PROSITE; PS0189; NTR; 1.
CC PROSITE; PS00288; TIMP; 1.
CC Metalloprotease inhibitor; Signal.
KW NON_TER
FT SIGNAL <1 2 BY SIMILARITY.
FT CHAIN 3 196 METALLOPROTEINASE INHIBITOR 2.
FT DOMAIN 3 128 NTR.
FT DISULFID 3 74 BY SIMILARITY.
FT DISULFID 5 103 BY SIMILARITY.
FT DISULFID 15 128 BY SIMILARITY.
FT DISULFID 130 177 BY SIMILARITY.
FT DISULFID 135 140 BY SIMILARITY.
FT DISULFID 148 169 BY SIMILARITY.
SQ SEQUENCE 196 AA; 21941 MW; 2ADAB9B54AF75381 CRC64;

Query Match 43.9%; Score 36; DB 1; Length 196;
Best Local Similarity 63.6%; Pred. No; 40;
Matches 7; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Oy 1 CANYASRGIRP 11
Db 177 CANY--RGAAP 185

Search completed: February 25, 2004, 06:42:59
Job time : 13 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 25, 2004, 06:40:25 ; Search time 39 Seconds

(without alignments)  
121.353 Million cell updates/sec

Title: US-09-700-643A-7

Perfect score: 82

Sequence: 1 CAWYASRGIRPVGRX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

SPTREMBL 25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	82.9	98	Q8WN12	Q8wn12 ovis aries
2	58	70.7	117	Q9W624	Q9w624 carassius a
3	45.5	55.5	420	Q93LZ7	Q93lz7 streptomyce
4	45	54.9	515	Q8VUB9	Q8vub9 bradyrhizob
5	44	53.7	141	Q8PJ39	Q8pj39 xanthomonas
6	44	53.7	146	Q8F7I1	Q8f7i1 xanthomonas
7	44	53.7	302	Q7U207	Q7u207 mycobacteri
8	44	53.7	315	P96274	P96274 mycobacteri
9	43	52.4	273	Q8ZTH7	Q8zth7 pyrobaculum
10	43	52.4	326	Q9F642	Q9f642 stigmatella
11	43	52.4	391	Q8I196	Q8i196 pseudomonas
12	43	52.4	402	Q82IE7	Q82ie7 streptomyce
13	43	52.4	403	Q8MRW9	Q8mrw9 drosophila
14	43	52.4	592	Q825Q8	Q825q8 streptomyce
15	43	52.4	1353	Q9V9Q7	Q9v9q7 drosophila
16	42	51.2	105	Q8UG50	Q8ug50 agrobacteri

17	42	51.2	113	16	Q8YL78	Q8yl78 anabaena sp
18	42	51.2	132	16	Q8PGJ6	Q8pgj6 xylella fas
19	42	51.2	132	16	Q87EQ6	Q87eq6 xylella fas
20	42	51.2	223	2	Q8RTA3	Q8rtc3 streptococc
21	42	51.2	250	16	Q8CWR1	Q8cwr1 streptococc
22	42	51.2	269	16	Q8X4D5	Q8x4d5 streptococc
23	42	51.2	288	16	Q82FH4	Q82fh4 streptomyce
24	42	51.2	289	16	Q89NJ5	Q89nj5 bradyrhizob
25	42	51.2	487	10	Q8LMS1	Q8lm51 arabidopsis
26	42	51.2	645	10	Q8L7I3	Q8l7i3 arabidopsis
27	42	51.2	647	6	Q95N04	Q95n04 sus scrofa
28	42	51.2	656	10	Q9WAM8	Q9nam8 arabidopsis
29	42	51.2	918	10	Q8SA86	Q8sa86 zea mays (m
30	42	51.2	1120	16	Q8UPG7	Q7upg7 rhodopirell
31	42	51.2	1120	16	Q857J2	Q857j2 mycobacteri
32	41	50.0	104	9	Q858M4	Q858m4 versinia pe
33	41	50.0	151	9	Q8FS67	Q8fs67 corynebacte
34	41	50.0	246	16	Q82VE9	Q82ve9 rhizobium m
35	41	50.0	256	11	Q8BUL8	Q8bul8 mus musculu
36	41	50.0	259	11	Q8NS20	Q8ns20 corynebacte
37	41	50.0	269	16	Q8NS20	Q8ns20 corynebacte
38	41	50.0	276	16	Q83929	Q83929 mycobacteri
39	41	50.0	276	16	Q7TZO8	Q7tzo8 mycobacteri
40	41	50.0	546	16	Q8YNR3	Q8ynr3 anabaena sp
41	41	50.0	966	2	Q47673	Q47673 escherichia
42	41	50.0	2368	2	Q93TW6	Q93tw6 stigmatella
43	40.5	49.4	115	16	Q7UVM9	Q7uvm9 rhodopirell
44	40.5	49.4	313	16	Q82YZ8	Q82yz8 rhizobium m
45	40	48.8	58	13	Q805E4	Q805e4 periophtal

## ALIGNMENTS

RESULT 1  
Q8WN12 PRELIMINARY; PRT; 98 AA.

AC Q8WN12; 2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Preprolactin-releasing peptide.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
[1]

RP SEQUENCE FROM N.A.  
RA Curlewis J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;  
RT "Prolactin-releasing peptide (PRP) in the ewe: cDNA cloning, mRNA  
RT distribution and effects on prolactin secretion in vitro and in  
RT vivo."  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF450453; AAL47178.1; -  
SQ SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;

Query Match 82.9%; Score 68; DB 6; Length 98;  
Best Local Similarity 92.3%; Pred. No. 0.0004;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14  
Db 40 AWYASRGIRPVGR 52

RESULT 2  
Q9W624 PRELIMINARY; PRT; 117 AA.

AC Q9W624;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE C-RF amide.  
 OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Carassius.  
 OX NCBI\_TaxID=7957;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Satake H., Minakata H., Fujimoto M.;  
 RT "Carassius RFamide (C-RF amide).";  
 RL Submitted (NGV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB020024; BAA76662.1; -;  
 SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C290 CRC64;  
 Query Match 70.7%; Score 58; DB 13; Length 117;  
 Best Local Similarity 66.7%; Pred. No. 0.026;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 WYASRGIRPVGR 14  
 |||:|:|:|:|:  
 DB 63 WYVGRGVPRIGR 74  
 RESULT 3  
 Q93LZ7 PRELIMINARY; PRT; 420 AA.  
 ID Q93LZ7  
 AC Q93LZ7  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Chain length factor-like protein.  
 GN AUR2B.  
 OS Streptomyces aureofaciens.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1894;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CM3239;  
 RA Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;  
 RT "Cloning and characterization of a new polyketide gene cluster in  
 Streptomyces aureofaciens CM3239.";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY033994; AAK61719.1; -;  
 DR HSSP; P39435; 1B3N.  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.  
 DR InterPro; IPR000794; Ketoacyl\_synth.  
 DR Pfam; PF00109; ketoacyl-synt; 1.  
 DR Pfam; PF02801; ketoacyl-synt\_C; 1.  
 SQ SEQUENCE 420 AA; 43011 MW; 3C27E22BE8C2DEA CRC64;  
 Query Match 55.5%; Score 45.5; DB 2; Length 420;  
 Best Local Similarity 55.6%; Pred. No. 16;  
 Matches 10; Conservative 2; Mismatches 1; Indels 5; Gaps 1;  
 QY 2 AWYASRGIRPVGR 14  
 |||:|:|:|:|:  
 DB 37 AWAAVLRGSGIRPVGR 54  
 RESULT 4  
 Q8VUB9 PRELIMINARY; PRT; 515 AA.  
 ID Q8VUB9  
 AC Q8VUB9  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Chlorin reductase subunit Y.  
 GN BCHY.  
 OS Bradyrhizobium sp. ORS278.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bradyrhizobiaceae; Bradyrhizobium.  
 OX NCBI\_TaxID=114615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ORS278;  
 RX MEDLINE=20570559; PubMed=11114184;  
 RA Giraud E., Hannibal L., Fardoux J., Vermeglio A., Dreyfus B.;  
 RT "Effect of Bradyrhizobium photosynthesis on stem nodulation of  
 RT Aeschynomene sensitiva.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14795-14800(2000).  
 DR EMBL; AF182374; AAL68697.1; -;  
 SQ SEQUENCE 515 AA; 54296 MW; 9170EB108F6A547F CRC64;  
 Query Match 54.9%; Score 45; DB 2; Length 515;  
 Best Local Similarity 70.0%; Pred. No. 24;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CAWYASRGIR 10  
 |||:|:|:|:  
 DB 382 CAWLESRGVR 391  
 RESULT 5  
 Q8PUJ39 PRELIMINARY; PRT; 141 AA.  
 ID Q8PUJ39  
 AC Q8PUJ39  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Protein-export membrane protein.  
 GN SEC OR XAC2706.  
 OS Xanthomonas axonopodis (pv. citri).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OX NCBI\_TaxID=92829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=306 / ATCC 13902 / XV 101;  
 RX MEDLINE=2202145; PubMed=12024217;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Aguiar R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Cipina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.B., Cursino-Santos J.R., El-Dorri H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.F.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,  
 RA Setubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities";  
 RL Nature 417:459-463(2002).  
 DR EMBL; AB011910; AAM37551.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0015450; F:protein translocase activity; IEA.  
 DR GO; GO:0009306; P:protein secretion; IEA.  
 DR InterPro; IPR004692; SecG.  
 DR Pfam; PF03840; SecG; 1.  
 DR PRINTS; PR01851; SECSEXPORT.  
 DR Complete proteome.  
 SQ SEQUENCE 141 AA; 13864 MW; CC8009663B05F5CA CRC64;  
 Query Match 53.7%; Score 44; DB 16; Length 141;  
 Best Local Similarity 70.0%; Pred. No. 8.8;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 AWYASRGIRP 11  
 |||:|:|:|:  
 DB 2 AWYASRGIRP 11

```

Db      56 AWYATHGARP 65

RESULT 6
QBP7T1
ID      QBP7T1      PRELIMINARY;      PRT;      146 AA.
AC      QBP7T1
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Protein-export membrane protein.
GN      SEG OR KC2530.
OS      Xanthomonas campestris (pv. campestris).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC      Xanthomonadaceae; Xanthomonas.
OX      NCBI_TaxID=340;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 33913 / NCPPB 528;
RX      MEDLINE=2022145; PubMed=12024217;
RA      da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA      Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA      Alves L.M.C., do Amaral A.W., Bertolini M.C., Camargo L.B.A., L.P.,
RA      Camarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
RA      Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA      Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA      Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA      Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA      Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA      Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA      Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA      Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA      Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA      Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA      Setubal J.C., Kitajima J.P.
RT      "Comparison of the genomes of two Xanthomonas pathogens with differing
RT      host specificities."
RL      Nature 417:459-463(2002).
DR      EMBL; AE012364; AA041802.1; --
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0015450; P:protein translocase activity; IEA.
DR      GO; GO:0009305; P:protein secretion; IEA.
DR      InterPro; IPR004692; SecG.
DR      Pfam; PF03840; SecG; I.
DR      PRINTS; PR01651; SECSEXPORT.
KW      Complete proteome.
SQ      SEQUENCE 146 AA; 14481 MW; 503ABAD982C04CB CRC64;

Query Match      53.7%; Score 44; DB 16; Length 146;
Best Local Similarity 70.0%; Pred. No. 9.1; Mismatches 2; Indels 0; Gaps 0;
Matches 7; Conservative 1;

QY      2 AWYASRGIRP 11
Db      56 AWYATHGARP 65

RESULT 7
Q7U207
ID      Q7U207      PRELIMINARY;      PRT;      302 AA.
AC      Q7U207
DT      01-OCT-2003 (TrEMBLrel. 25, Created)
DT      01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein.
GN      MB0436C.
OS      Mycobacterium bovis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1765;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=AF2122/97;
RX      MEDLINE=22709107; PubMed=12788972;
RA      Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA      Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA      Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA      Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT      "The complete genome sequence of Mycobacterium bovis."
RL      Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR      EMBL; BX248335; CAD93299.1; --
KW      Complete proteome.
SQ      SEQUENCE 302 AA; 32592 MW; 2A0EF63DA3B50D77 CRC64;

Query Match      53.7%; Score 44; DB 16; Length 302;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 AWYASRGIRP 11
Db      126 AWYASRDLP 135

RESULT 8
P96274
ID      P96274      PRELIMINARY;      PRT;      315 AA.
AC      P96274
DT      01-MAY-1997 (TrEMBLrel. 03, Created)
DT      01-MAY-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Hypothetical protein RV0428C.
GN      RV0428C OR MT0443 OR MTCY22G10.25C.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=H37Rv;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA      Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA      Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA      Hornsby T., Jagels K., Kitch A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA      Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence."
RL      Nature 393:537-544(1998).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CDC 1551 / Oshkosh;
RA      Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA      Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA      Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M., Salzberg S.L.,
RA      Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA      Bishai W.;
RT      "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT      laboratory strains."
RL      Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; Z84724; CAB06568.1; ALT_INIT.
DR      EMBL; AE006947; AA044666.1; --
DR      PIR; B70631; B70631.
DR      TIGR; MT0443; --
DR      TubercuList; RV0428c; --
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 315 AA; 34099 MW; 3BBE6399C4C0C7B7 CRC64;

Query Match      53.7%; Score 44; DB 16; Length 315;
Best Local Similarity 70.0%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 AWYASRGIRP 11

```

Db 139 AWYASRDLPQ 148

RESULT 9

Q8ZTH7 PRELIMINARY; PRT; 273 AA.

AC Q8ZTH7; 2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Morphine 6-dehydrogenase, conjectural.

GN PAE3247.

OS Pyrobaculum aerophilum.

OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;

OC Thermoproteaceae; Pyrobaculum.

OX NCBI\_TaxID=13773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=IM2 / ATCC 51768 / DSM 7523;

EX MEDLINE=21664397; PubMed=11792859;

RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,

RA Miller J.H.;

RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum

RT aerophilum";

RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).

DR EMBL: AE009920; AAL64784.1; -.

DR InterPro: IPR001395; Aldo/ket\_red.

DR Pfam: PF00248; aldo\_ket\_red; 1.

DR PRINTS: PR00069; ALDKETREDTASE.

DR ProDom: PD000288; Aldo/ket\_red; 1.

KW Complete proteome.

SQ SEQUENCE 273 AA; 30203 MW; 3F9D1208FB10DF7F CRC64;

Query Match 52.4%; Score 43; DB 17; Length 273;

Best Local Similarity 54.5%; Pred. No. 27;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 AWYASRGIRPV 12

Db 228 AWYKRGVWPI 238

RESULT 10

Q9F642 PRELIMINARY; PRT; 326 AA.

AC Q9F642;

DT 01-MAR-2001 (TREMBlrel. 15, Created)

DT 01-MAR-2001 (TREMBlrel. 15, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE MxCA.

GN MxCA.

OS Stigmatella aurantiaca.

OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;

OC Cythobacterineae; Cythobacteraceae; Stigmatella.

OX NCBI\_TaxID=41;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Sg a15;

EX MEDLINE=20485524; PubMed=11029592;

RA Silakowski B., Kunze B., Nordsiek G., Blocker H., Hofle G., Muller R.;

RT "The myxochelin iron transport regulon of the myxobacterium

RT Stigmatella aurantiaca Sg a15";

RL Eur. J. Biochem. 267:6476-6485(2000).

DR EMBL: AF299336; AAG31124.1; -.

DR InterPro: IPR001395; Aldo/ket\_red.

DR Pfam: PF00248; aldo\_ket\_red; 1.

DR ProDom: PD000288; Aldo/ket\_red; 1.

SQ SEQUENCE 326 AA; 35086 MW; 5C7979492B161DFF CRC64;

Query Match 52.4%; Score 43; DB 2; Length 326;

Best Local Similarity 70.0%; Pred. No. 32;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 AWYASRGIRPV 13

Db 85 AWFAEDGLKPG 96

RESULT 12

Q82IE7 PRELIMINARY; PRT; 402 AA.

AC Q82IE7;

DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Putative exoribonuclease large subunit.

GN XSEA OR SAV3211.

OS Streptomyces avermitilis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI\_TaxID=33903;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

EX MEDLINE=21477403; PubMed=11572948;

RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,

RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;

RT "Genome sequence of an industrial microorganism Streptomyces

RT avermitilis: deducing the ability of producing secondary

RT metabolites";

RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

RN [2]

RP SEQUENCE FROM N.A.

QY 2 AWYASRGIRP 11

Db 79 SWLASRGIKP 88

RESULT 11

Q88I96 PRELIMINARY; PRT; 391 AA.

AC Q88I96;

DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Conserved hypothetical protein.

GN PP3103.

OS Pseudomonas putida (strain KT2440).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI\_TaxID=160488;

RN [1]

RP SEQUENCE FROM N.A.

EX MEDLINE=22423060; PubMed=12534463;

RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,

RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,

RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolchay J.,

RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,

RA Chris Lee P., Holtzapfle B., Scanlan D., Tran K., Moazzez A.,

RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,

RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,

RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,

RA Fraser C.M.;

RT "Complete genome sequence and comparative analysis of the

RT metabolically versatile Pseudomonas putida KT2440.";

RL Environ. Microbiol. 4:799-808(2002).

DR EMBL: AE016785; AAN68711.1; -.

DR TIGR: PP3103; -.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 391 AA; 42589 MW; 840FACC07036F3A8 CRC64;

Query Match 52.4%; Score 43; DB 16; Length 391;

Best Local Similarity 50.0%; Pred. No. 39;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AWYASRGIRPV 13

Db 85 AWFAEDGLKPG 96

RESULT 12

Q82IE7 PRELIMINARY; PRT; 402 AA.

AC Q82IE7;

DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Putative exoribonuclease large subunit.

GN XSEA OR SAV3211.

OS Streptomyces avermitilis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI\_TaxID=33903;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

EX MEDLINE=21477403; PubMed=11572948;

RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,

RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;

RT "Genome sequence of an industrial microorganism Streptomyces

RT avermitilis: deducing the ability of producing secondary

RT metabolites";

RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

RN [2]

RP SEQUENCE FROM N.A.

```

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005034; BAC70922.1; -.
DR GO; GO:0009318; C:exodeoxyribonuclease VII complex; IEA.
DR GO; GO:0008855; F:exodeoxyribonuclease VII activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006308; P:DNA catabolism; IEA.
DR InterPro; IPR003753; Exonuc_VII_L.
DR Pfam; PF02601; Exonuc_VII_L; 1.
DR Pfam; PF01336; tRNA anti_1; 1.
DR TIGRFAMs; TIGR00237; xseA; 1.
KW Complete proteome.
SQ SEQUENCE 402 AA; 43844 MW; 863AB19C4C8F4A07 CRC64;

Query Match 52.4%; Score 43; DB 16; Length 402;
Best Local Similarity 47.4%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 1; Indels 8; Gaps 1;

Qy 3 WYASRG-----IRPVG 13
Db 91 WYAPRGQLSLRAEIRPIG 109

RESULT 13
Q8MRW9
ID Q8MRW9 PRELIMINARY; PRT; 403 AA.
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SD16815p.
GN CG31619 OR CG2122 OR CG2131 OR CG13236.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY119222; XAMS1082.1; -.
DR FlyBase; FBgn0051619; CG31619.
DR InterPro; IPR000884; TSPI.
DR PROSITE; PS50092; TSPI; 2.
SQ SEQUENCE 403 AA; 45343 MW; 1040B5951E8380C9 CRC64;

Query Match 52.4%; Score 43; DB 5; Length 403;
Best Local Similarity 61.5%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

Qy 1 CAWYASRGIRPVG 13
Db 322 CWYGSRR-RPAG 332

RESULT 14
Q825Q8
ID Q825Q8 PRELIMINARY; PRT; 592 AA.
AC Q825Q8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative 1,4-alpha-glucan branching enzyme.
GN SAV7399.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis; deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005050; BAC75110.1; -.
DR GO; GO:0003844; F:1,4-alpha-glucan branching enzyme activity; IEA.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0005378; P:glycogen biosynthesis; IEA.
DR InterPro; IPR006047; Alpha_amyl_cat.
DR InterPro; IPR006589; Alp_amyl_cat_sub.
DR InterPro; IPR006407; GLGB.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
DR SMART; SM00842; Amy; 1.
DR TIGRFAMs; TIGR01515; branching_enzym; 1.
KW Complete proteome.
SQ SEQUENCE 592 AA; 66805 MW; 3A735E0360BCC51 CRC64;

Query Match 52.4%; Score 43; DB 16; Length 592;
Best Local Similarity 72.7%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPV 12
Db 103 AWMAARGRPV 113

RESULT 15
Q9V9Q7
ID Q9V9Q7 PRELIMINARY; PRT; 1353 AA.
AC Q9V9Q7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG31619 protein.
GN CG31619 OR CG2122 OR CG2131 OR CG13236.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

```

George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Wandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,  
 RA Balow R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of *Drosophila melanogaster*";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Celnikier S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin J., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacleb J., Paragov V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RA "Sequencing of *Drosophila melanogaster* genome";  
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RL [3]  
 RN SEQUENCE FROM N.A.  
 RP Mirza S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaninkar J.S., Prochnik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnikier S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RA "Annotation of *Drosophila melanogaster* genome";  
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RL [4]  
 RN SEQUENCE FROM N.A.  
 RP Adams M.D., Celnikier S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RL [5]  
 RN SEQUENCE FROM N.A.  
 RP FlyBase;  
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AE003781; AA57229.2; -.

DR FlyBase; FBgn0051619; CG31619.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003598; IG C2.  
 DR InterPro; IPR000884; TSF1.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00090; tsf; 1; 8.  
 DR SMART; SM00409; IG; 1.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00209; TSPI; 11.  
 DR PROSITE; PS0835; IG LIKE; 1.  
 DR PROSITE; PS0092; TSF1; 8.  
 KW Immunoglobulin domain.  
 SQ SEQUENCE 1353 AA; 150210 MW; F83CDB090964272F CRC64;  
 Query Match 52.4%; Score 43; DB 5; Length 1353;  
 Best Local Similarity 61.5%; Pred. No. 1.5e+02;  
 Matches 8; Conservative 0; Mismatches 3; Indels 2; Gaps 1;  
 QY 1 CAVYASRGIRPVG 13  
 Db 1272 CWTYGSRR--RPAG 1282  
 RESULT 16  
 Q8UG50 PRELIMINARY; PRT; 105 AA.  
 ID Q8UG50  
 AC Q8UG50;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein Atul191.  
 GN Atul191 OR AGR.C.2196.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 OX NCBI\_TaxID=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Rao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RA "The genome of the natural genetic engineer *Agrobacterium tumefaciens*  
 C58";  
 RT Science 294:2317-2323(2001).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=21608551; PubMed=11743194;  
 RX Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Houello B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Quimio K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
*Agrobacterium tumefaciens* C58";  
 RL Science 294:2323-2328(2001).  
 DR EMBL; AE009082; AAL42203.1; -.  
 DR EMBL; AE008047; AAK86984.1; -.  
 DR PIR; A57505; A97505.  
 DR PIR; A57233; AE7233.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 105 AA; 11749 MW; 0053AEA20849A1D0 CRC64;  
 Query Match 51.2%; Score 42; DB 16; Length 105;

Best Local Similarity 66.7%; Pred. NO. 14;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AWYASRGIR 10  
| | | | |  
Db 54 AWYARGVQ 62

RESULT 17  
Q8YL78 PRELIMINARY; PRT; 113 AA.  
AC Q8YL78;  
DT 01-VAR-2002 (TrEMBLrel. 20, Created)  
DT 01-VAR-2002 (TrEMBLrel. 20, Last sequence update)  
DE Hypothetical protein Alr7057.  
GN ALR7057.  
OS Anabaena sp. (strain PCC 7120).  
OG Plasmid pCC7120alpha.  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
OX NCBI\_TaxID=103690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21595285; PubMed=11759840;  
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.,  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
RT cyanobacterium Anabaena sp. strain PCC 7120.";  
RL DNA Res. 8:205-213(2001).  
DR EMBL; AF003600; BAB78141.1; -;  
DR PIR; A2484; A12484.  
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
KW Plasmid; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 113 AA; 13058 MW; CF5ED411086C06A3 CRC64;

Query Match 51.2%; Score 42; DB 16; Length 113;  
Best Local Similarity 70.0%; Pred. NO. 15;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AWYASRGIRP 11  
| | | | |  
Db 64 AWYAKGIEP 73

RESULT 18  
Q9PGJ6 PRELIMINARY; PRT; 132 AA.  
AC Q9PGJ6;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Protein-export membrane protein.  
GN XF0304.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9a5G;  
RX STRAIN=9a5G;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu P.A., Acencio M.,  
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.H.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
RA Paciniani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,  
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
RL Nature 406:151-159(2000).  
DR EMBL; AB003883; AAF83115.1; -;  
DR PIR; B82824; B82824.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0015450; P:protein translocase activity; IEA.  
DR GO; GO:0009306; P:protein secretion; IEA.  
DR InterPro; IPR004692; SecG.  
DR Pfam; PF03840; SecG; 1.  
DR PRINTS; PR01651; SECSEXPORT.  
DR TIGRPFAMS; TIGR00810; secG; 1.  
KW Complete proteome.  
SQ SEQUENCE 132 AA; 13530 MW; 757240DD2D7A1308 CRC64;

Query Match 51.2%; Score 42; DB 16; Length 132;  
Best Local Similarity 61.5%; Pred. NO. 18;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14  
| | | | |  
Db 70 AWYASRGHVAQ 82

RESULT 19  
Q87EQ6 PRELIMINARY; PRT; 132 AA.  
AC Q87EQ6;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Protein-export membrane protein SecG.  
GN SECGR ED0246.  
OS Xylella fastidiosa (strain Temecul / ATCC 700964).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xylella.  
OX NCBI\_TaxID=183190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22421331; PubMed=12533478;  
RX Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B., Moon D.H.,  
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,  
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,  
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorry H., Tsai S.M.,  
RA Carrier H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,  
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,  
RA Marino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,  
RA Baia G.S., Bianco S.R., Brito M.S., Cannava F.S., Celestino A.V.,  
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,  
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,  
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,  
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,  
RA Kitajima J.P.;  
RT "Comparative analyses of the complete genome sequences of Pierce's  
RT disease and citrus variegated chlorosis strains of Xylella  
RT fastidiosa.";  
RL J. Bacteriol. 185:1018-1026(2003).

DR EMBL; AE012554; AAC28133.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0015450; F:protein translocase activity; IEA.  
 DR GO; GO:0009306; P:protein secretion; IEA.  
 DR InterPro; IPR004692; SecG.  
 DR Pfam; PF03840; SecG; 1.  
 DR PRINTS; PR01651; SECSEXPORT.  
 DR Complete proteome.  
 SQ SEQUENCE 132 AA; 13605 MW; 676040C62CC77808 CRC64;

Query Match 51.2%; Score 42; DB 16; Length 132;  
 Best Local Similarity 61.5%; Pred. No. 18;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AWYASRGIRPVGR 14  
 ||||| :  
 Db 70 AWYASRGHHSVAQ 82

RESULT 20  
 Q8RTA3 PRELIMINARY; PRT; 223 AA.  
 AC Q8RTA3;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE LicD2 (Fragment).  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WB44.1;  
 RA Weiser J.N.; Bae D.H.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF467751; AAL77432.1; -;  
 DR InterPro; IPR007074; LicD.  
 DR Pfam; PF04991; LicD; 1.  
 FT NON TER 223 223  
 SQ SEQUENCE 223 AA; 26566 MW; C0FC9A3A7AC75384 CRC64;

Query Match 51.2%; Score 42; DB 2; Length 223;  
 Best Local Similarity 54.5%; Pred. No. 32;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAWYASRGIRP 11  
 :||| :  
 Db 166 CSWYALRFVNP 176

RESULT 21  
 Q912F2 PRELIMINARY; PRT; 250 AA.  
 AC Q912F2;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein PA1952.  
 GN PA1952.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K.; Pham X.-O.T.; Erwin A.L.; Mizoguchi S.D.; Warrenner P.;  
 RA Hickey M.J.; Brinkman F.S.L.; Hufnagle W.O.; Kowalik D.J.; Lagrou M.;  
 RA Garber R.L.; Goltry L.; Tolentino E.; Westbrock-Wadman S.; Yuan Y.;  
 RA Brody L.L.; Coulter S.N.; Folger K.K.; Kas A.; Larbig K.; Lim R.M.;  
 RA Smith K.A.; Spencer D.H.; Wong G.K.-S.; Wu Z.; Paulsen I.T.;

RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an  
 RT opportunistic pathogen."  
 RL Nature 406:959-964 (2000).  
 DR EMBL; AE004622; AAG05340.1; -;  
 DR PIR; G83400; G83400.  
 DR InterPro; IPR000437; Prok\_lipoprot.S.  
 DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 250 AA; 25619 MW; B97F6BE28D792C2 CRC64;

Query Match 51.2%; Score 42; DB 16; Length 250;  
 Best Local Similarity 53.8%; Pred. No. 36;  
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CAWYASRGIRPVG 13  
 ||| :||| :  
 Db 237 CAWEQLRALRPSG 249

RESULT 22  
 Q8CWR1 PRELIMINARY; PRT; 269 AA.  
 ID Q8CWR1;  
 AC Q8CWR1;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE LicD protein.  
 GN LICD2 OR SPRI152.  
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=171101;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21429245; PubMed=11544234;  
 RA Hoskins J.; Albom W.E. Jr.; Arnold J.; Blaszczyk L.C.; Burett S.;  
 RA DeHoff B.S.; Estrem S.T.; Fritz L.; Fu D.-J.; Fuller W.; Geringer C.;  
 RA Gilmore R.; Glass J.S.; Khoja H.; Kraft A.R.; Lagace R.E.;  
 RA LeBlanc D.J.; Lee L.N.; Lefkowitz E.J.; Lu J.; Matsushima P.;  
 RA McAhren S.M.; McHenry M.; McLeaster K.; Mundy C.W.; Nicas T.I.;  
 RA Norris F.H.; O'Gara M.; Peery R.B.; Robertson G.T.; Rocky P.;  
 RA Sun P.-M.; Winkler M.E.; Yang Y.; Young-Bellido M.; Zhao G.;  
 RA Zook C.A.; Baltz R.H.; Jaskunas S.R.; Rostek P.R. Jr.; Skatrud P.L.;  
 RA Glass J.I.;  
 RT "Genome of the bacterium *Streptococcus pneumoniae* strain R6.";  
 RL J. Bacteriol. 183:5709-5717 (2001).  
 DR EMBL; AE008487; AAK99955.1; -;  
 DR PIR; A95148; A95148.  
 DR PIR; G98015; G98015.  
 DR InterPro; IPR007074; LicD.  
 DR Pfam; PF04991; LicD; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 269 AA; 32100 MW; 31A152DFAA480A10 CRC64;

Query Match 51.2%; Score 42; DB 16; Length 269;  
 Best Local Similarity 54.5%; Pred. No. 39;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAWYASRGIRP 11  
 :||| :  
 Db 166 CSWYALRFVNP 176

RESULT 23  
 Q9X4D5 PRELIMINARY; PRT; 269 AA.  
 ID Q9X4D5;  
 AC Q9X4D5;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE LICD2.  
 GN LICD2 OR SP1274.





SEQUENCE FROM N.A.  
 RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,  
 RA Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu S., Mukharsky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC F2E2 from chromosome  
 I";  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,  
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu S., Mukharsky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC069252; AAF86559.1;  
 SQ SEQUENCE 487 AA; 55018 MW; DC46ED10FA5AD679 CRC64;

Query Match 51.2%; Score 42; DB 10; Length 487;  
 Best Local Similarity 54.5%; Pred. No. 74;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CAWYASRGIRPV 11  
 Db 126 CAWYLSNHVEP 136

RESULT 27  
 O81713 PRELIMINARY; PRT; 645 AA.  
 AC O81713  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Late elongated hypocotyl protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=98319236; PubMed=9657154;  
 RA Schaffer R., Raney N., Samach A., Corden S., Putterill J.,  
 RA Carre I.A., Coupland G.;  
 RT "The late elongated hypocotyl mutation of Arabidopsis disrupts  
 RT circadian rhythms and the photoperiodic control of flowering.";  
 RL Cell 93:1219-1229 (1998).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.  
 DR EMBL: AJ006404; CAA07004.1;  
 DR TRANSFAC: T02870;  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR InterPro: IPR001005; Myb DNA binding.  
 DR InterPro: IPR006447; Myb SHACKYF.  
 DR Pfam: PF00249; myb DNA-binding; 1.  
 DR SMART: SMC0717; SANT; 1.  
 DR TIGRFAMs: TIGR01557; myb\_SHACKYF; 1.

DR PROSITE: PS50090; MYB\_3; 1.  
 KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 645 AA; 70425 MW; D184A30CCFB3EA77 CRC64;  
 Query Match 51.2%; Score 42; DB 10; Length 645;  
 Best Local Similarity 63.6%; Pred. No. 1e+02; 2; Indels 0; Gaps 0;  
 Matches 7; Conservative 2; Mismatches 2;  
 Qy 2 AWYASRGIRPV 12  
 Db 357 AWWASHGLLPV 367

## RESULT 28

Q95N04 PRELIMINARY; PRT; 647 AA.  
 AC Q95N04  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Dihydrolipoamide acetyltransferase precursor (EC 2.3.1.12).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Koike K.;  
 RT "Isolation and characterization of the cDNA encoding the  
 RT dihydrolipoamide acetyltransferase component of the porcine pyruvate  
 RT dehydrogenase complex";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 LIPOYL-BINDING DOMAINS.  
 DR EMBL: AB036739; BAB61720.1;  
 DR GO: GO:0005739; C:mitochondrion; IEA.  
 DR GO: GO:0045254; C:pyruvate dehydrogenase complex; IEA.  
 DR GO: GO:0008415; F:acetyltransferase activity; IEA.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0004742; F:dihydrolipoamide S-acetyltransferase activity; IEA.  
 DR GO: GO:0005515; F:protein binding; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0004812; F:tRNA ligase activity; IEA.  
 DR GO: GO:0006418; P:amino acid activation; IEA.  
 DR GO: GO:0006096; P:glycolysis; IEA.  
 DR GO: GO:0008152; P:metabolism; IEA.  
 DR InterPro: IPR001078; 2Oxoacid dh.  
 DR InterPro: IPR006257; AceF long.  
 DR InterPro: IPR000089; Biotin lipoyl.  
 DR InterPro: IPR004167; E3 binding.  
 DR InterPro: IPR003016; Lipoyl BS.  
 DR InterPro: IPR003412; tRNA-synt\_1.  
 DR Pfam: PF00198; 2-oxoacid dh; 1.  
 DR Pfam: PF00364; biotin\_lipoyl; 2.  
 DR Pfam: PF02817; e3\_binding; 1.  
 DR TIGRFAMs: TIGR01115; 2Oxoacid dh; 1.  
 DR TIGRFAMs: TIGR01349; PDHac trf mito; 1.  
 DR PROSITE: PS00178; AA TRNA\_LIGASE\_I; 1.  
 DR PROSITE: PS00189; LIPOYL\_2.  
 KW Acyltransferase; Lipoyl; Signal; Transferase; Mitochondrion.  
 FT SIGNAL 1 86  
 FT CHAIN 87 647  
 SQ SEQUENCE 647 AA; 69034 MW; 3BB554B0D5F8235E CRC64;

Query Match 51.2%; Score 42; DB 6; Length 647;

Best Local Similarity 50.0%; Pred. No. 1e+02; 6; Indels 0; Gaps 0;  
 Matches 7; Conservative 1; Mismatches 6;

Qy 1 CAWYASRGIRPVGR 14  
 Db 59 CGWSATRGATPQNR 72

```

RESULT 29
Q9WAM8      PRELIMINARY;      PRT;      656 AA.
ID Q9WAM8;
AC Q9WAM8;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE T25K16.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RA Dunn P., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Federspiel N.A.,
RA Theologis A., Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC T25K16 from chromosome
RT I.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL; AC007323; AAF26474.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001005; MYB DNA binding.
DR InterPro; IPR006447; MYB SHAQKYF.
DR Pfam; PF00249; MYB DNA-binding; 1.
DR SMART; SM00717; SANT; 1.
DR TIGRFAMs; TIGR01557; myb_SHAQKYF; 1.
DR PROSITE; PS5090; MYB_3; 1.
DR DNA-binding; Nuclear protein.
KW SEQUENCE 656 AA; 71794 MW; E3D08CF31F739A2E CRC64;

Query Match      51.2%; Score 42; DB 10; Length 656;
Best Local Similarity 63.6%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 AWYASRGIRPV 12
Db      368 AWWASHGLLPV 378

RESULT 30
Q8SA86      PRELIMINARY;      PRT;      918 AA.
ID Q8SA86;
AC Q8SA86;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Putative retrotransposon protein.
GN Z195D10.18.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae.
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1];
RP SEQUENCE FROM N.A.
RA STRAIN=cv. B73;
RA Ramakrishna W., Emberton J., SanMiguel P., Bennetzen J.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. B73;
RC Doebley J.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. B73;
RA Liaca V., Linton E.W., Young S., Kovchok S., Messing J.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF466646; AAL76006.1; -
DR InterPro; IPR007321; Transposase 28.
DR Pfam; PF04195; Transposase 28; 1.
SQ SEQUENCE 918 AA; 103639 MW; 11F4E95C34039495 CRC64;

Query Match      51.2%; Score 42; DB 10; Length 918;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CAWVASRG 8
Db      836 CAWVASRG 843

RESULT 31
Q7UPG7      PRELIMINARY;      PRT;      1120 AA.
ID Q7UPG7;
AC Q7UPG7;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Rhodospirella baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Haitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL; BX294145; CAD75095.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1120 AA; 123221 MW; 67C757B73421BF9B CRC64;

Query Match      51.2%; Score 42; DB 16; Length 1120;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 AWYASRGIRP 11
Db      206 AWWAXGIRP 215

RESULT 32
Q857J2      PRELIMINARY;      PRT;      104 AA.
ID Q857J2;
AC Q857J2;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Gp30.
OS Mycobacteriophage Bx22.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=205870;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=22592660; PubMed=12705866;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,

```

RA Hatfull G.F.;  
RT "Origins of highly mosaic mycobacteriophage genomes.";  
RL Cell 113:171-182(2003).  
DR EMBL; AY129332; AAN01784.1; -.  
SQ SEQUENCE 104 AA; 11312 MW; C72FE2B938D1ED11 CRC64;

Query Match 50.0%; Score 41; DB 9; Length 104;  
Best Local Similarity 46.2%; Pred. No. 21;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
|||:|:|:|:  
DB 88 AWIIRNVTPVGK 100  
|||:|:|:|:

RESULT 33  
Q858M4  
ID Q858M4 PRELIMINARY; PRT; 151 AA.  
AC Q858M4;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Endonuclease.  
OS Versinia pestis phage phiA1122.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
OC T7-like viruses  
OX NCBI\_TaxID=227720;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Garcia E., Elliott J.M., Ramanculov E., Chain P.S., Chu M.C.,  
RA Molineux I.J.;  
RT "The genome sequence of Versinia pestis bacteriophage PhiA1122 reveals  
RT an intimate history with the coliphage T3.";  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY247822; AAP20511.1; -.  
DR GO; GO:0008833; F:deoxyribonuclease IV (phage-T4-induced) act. .; IEA.  
DR GO; GO:0004519; F:endonuclease activity; IEA.  
DR GO; GO:0015074; P:DNA integration; IEA.  
DR GO; GO:0016032; P:viral life cycle; IEA.  
DR InterPro: IPR008029; Phage\_endo\_I.  
DR Pfam: PF05367; Phage\_endo\_I; 1.  
KW Endonuclease.  
SQ SEQUENCE 151 AA; 17402 MW; 1EA8D071D41BB979 CRC64;

Query Match 50.0%; Score 41; DB 9; Length 151;  
Best Local Similarity 80.0%; Pred. No. 31;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 YASRGIRPVG 13  
|||:|:|:|:  
DB 5 YAARGIRKVG 14  
|||:|:|:|:

RESULT 34  
Q8FS67  
ID Q8FS67 PRELIMINARY; PRT; 246 AA.  
AC Q8FS67;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative oxidoreductase.  
GN C80537.  
OS Corynebacterium efficiens.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=152794;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,  
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,  
RA Usuda Y., Sugimoto S.;  
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP005215; BAC17347.1; -.  
DR InterPro: IPR001395; Aldo/ket\_red.  
DR Pfam: PF00248; aldo\_ket\_red; 1.  
DR PRINTS; PR00069; ALDKETREDTASE.  
DR ProDom; PD000288; Aldo/ket\_red; 1.  
DR PROSITE; PS00798; ALDOKETO\_REDUCTASE\_1; 1.  
DR PROSITE; PS00062; ALDOKETO\_REDUCTASE\_2; 1.  
DR PROSITE; PS00063; ALDOKETO\_REDUCTASE\_3; 1.  
KW Complete proteome.  
SQ SEQUENCE 246 AA; 27766 MW; 9BB9A6402F9D27DD CRC64;

Query Match 50.0%; Score 41; DB 16; Length 246;  
Best Local Similarity 53.8%; Pred. No. 53;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AWYASRGIRPVGR 14  
|||:|:|:|:  
DB 187 AWHHARGIVPIPR 199  
|||:|:|:|:

RESULT 35  
Q92VE9  
ID Q92VE9 PRELIMINARY; PRT; 256 AA.  
AC Q92VE9;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein RB0757.  
GN RB0757 OR SM21253.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OG Plasmid pSymB (megaplasmid 2).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=213396508; PubMed=11481431;  
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,  
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Souzy J.,  
RA Golding B., Puchler A.;  
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-  
RT fixing endosymbiont Sinorhizobium meliloti.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).  
DR EMBL; AL603644; CAC49157.1; -.  
DR PIR; E95936; E95936.  
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
KW Plasmid; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 256 AA; 28459 MW; 54B054834CEC7C39 CRC64;

Query Match 50.0%; Score 41; DB 16; Length 256;  
Best Local Similarity 54.5%; Pred. No. 55;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 WYASRGIRPVG 13  
|||:|:|:|:  
DB 185 WHGTRGCRPYG 195  
|||:|:|:|:

RESULT 36  
Q8BUL8  
ID Q8BUL8 PRELIMINARY; PRT; 259 AA.  
AC Q8BUL8;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical pleckstrin homology.  
GN C920005C14RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;

```

RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=1246851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK083325; BAC38865.1; -.
DR MGD; MGI:2443609; C920005C14Rik.
DR InterPro; IPR001849; PH.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS00003; PH DOMAIN; 1.
KW Hypothetical protein.
SQ SEQUENCE 259 AA; 28734 MW; 2B27DFD35DFA0906 CRC64;

Query Match 50.0%; Score 41; DB 11; Length 259;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 RGIRPVGR 14
DB 188 RGIRPIGR 195

RESULT 37
Q8NSZ0 PRELIMINARY; PRT; 269 AA.
AC Q8NSZ0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Aldo/keto reductases, related to diketoglucuronate reductase (EC
DE 1.1.1.-).
GN CGL0525.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF005275; BAB97918.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; aldo ket red; 1.
DR PRINTS; PR00069; ALDKETRDTASE.
DR ProDom; PD00288; Aldo/ket_red; 1.
DR PROSITE; PS00798; ALDOKETO_REDUCTASE_1; 1.
DR PROSITE; PS00062; ALDOKETO_REDUCTASE_2; 1.
DR PROSITE; PS00063; ALDOKETO_REDUCTASE_3; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 269 AA; 30046 MW; F858A0130AE2891A CRC64;

Query Match 50.0%; Score 41; DB 16; Length 269;
Best Local Similarity 53.8%; Pred. No. 58;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AWVASRGIRPVGR 14
DB 210 AWHARGIVPIPR 222

RESULT 38
O53929 PRELIMINARY; PRT; 276 AA.
ID O53929
AC O53929;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)

```

```

DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Rv1716.
GN Rv1716 OR MT1755 OR MTV048.03 OR MTCY04C12.01.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL0222003; CAAL7613.1; -.
DR EMBL; AE007037; AAK46027.1; -.
DR PIR; D70817; D70817.
DR TIGR; MT1755; -.
DR TuberculList; Rv1716; -.
DR InterPro; IPR007325; Cyclase.
DR Pfam; PF04199; Cyclase; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 178 178 S->G (IN REF. 2).
FT CONFLICT 276 276 V->A (IN REF. 2).
SQ SEQUENCE 276 AA; 30073 MW; 8F13DC9972A808D8 CRC64;

Query Match 50.0%; Score 41; DB 16; Length 276;
Best Local Similarity 45.5%; Pred. No. 60;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 3 WYASRGIRPVG 13
DB 150 WFAAGVKVAVG 160

RESULT 39
Q7TZQ8 PRELIMINARY; PRT; 276 AA.
AC Q7TZQ8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved hypothetical protein.
GN MB1744.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,

```

RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
RA Parkhill J., Barrrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.,  
RT "The complete genome sequence of *Mycobacterium bovis*."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).  
DR EMBL; BX248340; CAD94447.1; -.  
KW Complete proteome.  
SQ SEQUENCE 276 AA; 30015 MW; 9107C098AEB5C904 CRC64;

Query Match 50.0%; Score 41; DB 16; Length 276;  
Best Local Similarity 45.5%; Pred. No. 60;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 3 WYASRGIRPVG 13  
|:|:|:|:|  
DB 150 WFAAGYKAVG 160

RESULT 40  
Q8YNR3  
ID Q8YNR3 PRELIMINARY; PRT; 546 AA.  
AC Q8YNR3;  
DT 01-MAR-2002 (TRENBLrel. 20, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Hypothetical protein All4499.  
GN All4499.  
OS Anabaena sp. (strain PCC 7120).  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
OX NCBI\_TaxID=103690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21595285; PubMed=11759840;  
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsunoto M., Matsuno A., Muraki A.,  
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
cyanobacterium *Anabaena* sp. strain PCC 7120.";  
RL DNA Res. 8:205-213(2001).  
DR EMBL; AP003596; BAB76198.1; -.  
DR PIR; AC2368; AC2368.  
DR InterPro; IPR001119; SLH.  
DR Pfam; PF00395; SLH; 1.  
DR PROSITE; PS01072; SLH\_DOMAIN; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 546 AA; 58563 MW; 7A6A8E3961F72316 CRC64;

Query Match 50.0%; Score 41; DB 16; Length 546;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 10; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 2 AWY--ASRGIRPVGR 14  
| | | | |  
DB 457 AWYWAASIGIRDFGR 471

Search completed: February 25, 2004, 06:43:51  
Job time : 40 secs